

DOCUMENT RESUME

ED 210 848

EC 140 648

AUTHOR Hicks, Doin E., Ed.; And Others
 TITLE Assessment of Hearing Impaired Youth.
 INSTITUTION Gallaudet Coll., Washington, D.C.
 PUB DATE 80
 NOTE 71p.; Photographs may not reproduce clearly.
 AVAILABLE FROM Directions, Box 5664, Washington, DC 20016 (\$10.00 per year, \$3.50 single copy).
 JOURNAL CIT Directions; v1 n4 1980
 EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
 DESCRIPTORS Audiology; Communication Skills; Criterion Referenced Tests; Educational Diagnosis; Elementary Secondary Education; *Evaluation Methods; *Hearing Impairments; Informal Assessment; Job Skills; Medical Evaluation; Psychoeducational Methods; Psychological Needs; *Student Evaluation

ABSTRACT

The issue of Directions contains 11 articles on assessment of hearing impaired individuals. Entries have the following titles and authors: "Classroom Assessment Techniques for Hearing Impaired Students--A Literature Review" (B. McKee, M. Hausknecht); "Informal Assessment of Hearing Impaired Students In the Classroom" (B. Culhane, R. Hein); "Educational Assessment for Placement" (R. Lennan); "Psychoeducational Diagnosis of the Hearing Impaired--A Student Monitoring System" (J. Heriot, G. Maddock); "Vocational Assessment--Approaches and Implications" (M. Vernon); "Assessing Communication Skills" (A. Spragins, D. Cckely); "Audiological Evaluation" (A. Maxon); "Medical Evaluation--An Overview" (R. Freeman); "Psychosocial Needs of Deaf College Students From a Faculty/Staff Perspective" (J. Covill-Servo, W. Garrison); "Assessing Labor Cost of Objectives" (S. Engelmann, A. Granzin); and "Using Criterion-Referenced Tests With Handicapped Students" (L. Anderson-Inman). (SB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED210848

Directions

Gallaudet College

Assessment of Hearing Impaired Youth

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Update on Academic, Professional
Career and Research Activities
Volume 1, Number 4, 1980
ISSN #0197-0224

"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFORM ONLY
HAS BEEN GRANTED BY

Charlotte
Franklin Shepard

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Contents

5

Foreword

7

Introduction

9

Classroom Assessment
Techniques for Hearing Impaired
Students: A Literature Review

Barbara G. McKee

Michael A. Hausknecht

16

Informal Assessment of
Hearing Impaired Students
In the Classroom

Barry R. Culhane

Ronald Hein

22

Educational Assessment
for Placement

Robert K. Lennan

26

Psychoeducational Diagnosis of
the Hearing Impaired: A Student
Monitoring System

James T. Heriot

Gerard J. Maddock

31

Vocational Assessment:
Approaches and Implications

McCay Vernon

36

Assessing Communication Skills

Anne B. Spragins

Dennis R. Cokely

39

Audiological Evaluation

Antonia B. Maxon

46

Medical Evaluation: An Overview

Roger D. Freeman

48

Psychosocial Needs of
Deaf College Students
From a Faculty/Staff Perspective

Jenny L. Covill-Servo

Wayne M. Garrison

54

Assessing Labor Cost of Objectives

Siegfried Engelmann

Alex Granzin

63

Using Criterion-Referenced Tests
With Handicapped Students

Lynne Anderson-Inman

About the Authors

Barbara G. McKee

Michael Hausknecht

Department of Educational

Research & Development

National Technical Institute

for the Deaf

Barry R. Culhane

Division of Human Studies
National Technical Institute
for the Deaf

Ronald Hein

Department of Educational
Research & Development
National Technical Institute
for the Deaf

Robert K. Lennan

Superintendent
California School for the Deaf
Riverside

James T. Heriot

Gerard J. Maddock

Arizona Diagnostic Treatment
& Education Center

A Division of Arizona State School
for the Deaf and Blind

McCay Vernon

Powrie V. Doctor Chair

Gallaudet College

Anne B. Spragins

Psychology Department

Gallaudet College

Dennis R. Cokely

Linguistics Lab

Gallaudet College

Antonia B. Maxon

Mainstream Project

College of Liberal Arts & Sciences

University of Connecticut

Roger D. Freeman

Psychiatry Department

University of British Columbia

Jenny L. Covill-Servo

Wayne M. Garrison

Department of Educational

Research & Development

National Technical Institute
for the Deaf

Siegfried Engelmann

Alex Granzin

College of Education

University of Oregon

Lynne Anderson-Inman

College of Education

University of Oregon

Directions

An Official Publication of
Gallaudet College

Dr. George E. Muth
Chairman of the Board
Dr. Edward C. Merrill, Jr.
President
Dr. Leonard M. Elstad
President Emeritus

Editorial Staff

Doin E. Hicks
Editor

Clarence M. Williams
Co-Editor

Charlotte Hawkins-Shepard
Special Editorial Assistant
Kathleen M. Clarke
Technical Editor

Gallaudet College Press Staff

Elaine Costello
Director

Leonard Lane
Associate Director
Donna Simons
Production Specialist

Editorial Consultant Board

Eugene Bergman
Gallaudet College
Gertrude Galloway
National Association of the Deaf
Mervin D. Garretson
Gallaudet College
Kenneth R. Lane
Fairfax, Virginia
Edgar L. Lowell
John Tracy Clinic
Kathryn P. Meadow
Gallaudet College
E. Ross Stuckless
National Technical Institute for the
Deaf

Editorial Consultants

Edna P. Adler
Rehabilitation Services
Administration, DHEW

Jack W. Birch
University of Pittsburgh

Richard G. Brill
California School for the Deaf
Riverside (Retired)

Elaine R. Costello
Gallaudet College
Sr. Frances Cronin
St. Joseph's School for the Deaf

Robert R. Davila
Gallaudet College

Gerilee Gustason
Gallaudet College

Ralph L. Hoag
Arizona School for the Deaf

Irving K. Jordan, Jr.
Gallaudet College

William J. A. Marshall
Gallaudet College

John K. Miller
University of Rochester

S. Richard Silverman
Central Institute for the Deaf
(Retired)

David A. Spidal
New York School for the Deaf
White Plains

Robert E. Stepp
University of Nebraska

Doin E. Hicks
Vice President for Research and
Professor, Educational
Administration, Gallaudet College

Clarence M. Williams
Coordinator of Faculty and
Graduate Student Research/
Division of Research; Professor,
Department of Education,
Gallaudet College

The Editorial Staff of *Directions* seeks to provide an open forum for the free expression of competent opinion relating to its area of interest. The materials printed herein are the expressions of the writers and not necessarily a statement of policy or a reflection of the opinion of the editors of *Directions* or of Gallaudet College.

The Gallaudet College Press publishes *Directions* four times per year. Subscription rates are \$10.00 per year in the United States, \$3.50 per single copy, and \$12.00 per year outside the U.S., \$4.00 per single copy. Subscription orders should be sent to *Directions*, Box 5664, Washington, D.C. 20016. Address changes should reach the subscription office by the first of the month preceding a change. Subscribers must notify the Post Office that they will guarantee forwarding postage for undelivered copies. Other claims for undelivered copies must be made within four months of publication.

Permission to quote a portion of any paper in this issue is hereby granted if permission is also obtained from the author(s) and if credit is given to the author(s) and this journal.

Copyright 1980 Gallaudet College.
All rights reserved.

Gallaudet College is an equal opportunity employer/educational institution. Programs and services offered by Gallaudet College receive substantial financial support from the Department of Education.

Purpose

Directions is a publication of Gallaudet College intended to provide professionals in the field of deafness with information perceived as important to the education and welfare of deaf persons. *Directions* is a vehicle for disseminating outstanding ideas, important papers, validated experience and research findings—primarily originating on Kendall Green—viewed to be of interest to the profession at large. It will update significant progress in academic, professional development, career and research areas. The goals of the publication are to: (1) foster the continuing development of theory and practice, and (2) encourage extensive and thorough discussion of problems related to and/or drawn from such innovation.

Information for Authors

Each issue of *Directions* will deal with a single topic identified as having interest to the profession at large. *Directions* issues generally are planned by editors and/or guest editors many months in advance and most of the articles are solicited from persons who have published original work on the topic. However, from time to time the journal will solicit manuscripts on a particular topic from individuals outside the College community, either directly or through a call for papers. Manuscripts and editorial correspondence should be sent to Doin E. Hicks, Editor, *Directions*, Gallaudet College, Washington, D.C., 20002.

Manuscripts. An original and two copies should be submitted for publication consideration. The manuscript should be typed—on 216 x 280mm heavy duty, white bond paper—double spaced, with margins of at least 38mm at top, bottom and

left, and 25mm at right. Pages should be numbered consecutively in the upper right-hand corner of each page. Page 1 of the original and one copy of the manuscript should include (1) the title of the article, (2) the names (without degrees) and institutional affiliations of all authors, and (3) the name, address, and telephone numbers of the author to whom editorial correspondence should be addressed.

Page 2 of the manuscript should contain an abstract not to exceed 200 words. The abstract should begin with reference to the subject of the article. The abstract should convey to the reader as much information as possible about highlights of the manuscript.

The text or body of the manuscript should begin on page 3. In this section, the liberal use of major headings and subheadings is recommended. First-level headings designate the major divisions of the paper. Usually no more than three heading levels should be used. Descriptive headings are recommended. Use of abbreviations, italics, and parenthetical statements is discouraged. The metric system should be used on all measurements. Referencing and other matters of bibliographic style should follow the form set forth in the Publication Manual of the American Psychological Association.

Citation of grant or contract support, change in affiliation of one or more authors, and other similar information, should be placed in the Acknowledgement.

Tables. Tables may be used where they will simplify presentation of information. Otherwise they should not be used. Each table should be typed, double spaced, on a separate sheet of paper. The tables should be numbered consecutively with Arabic numerals. The tables should be re-

ferred to in the text but the tables should be designed to be self-explanatory. Each table heading should be concise, yet should include a description of what the table contains.

Figures. Each manuscript copy should contain all figures. However, only one set of publication-quality figures are needed. The figures should be numbered consecutively with Arabic numerals. All figures should be referred to in the text, but each figure should be designed to be self-explanatory. Each figure should be placed on a separate page, without legend. All legends should be typed, double spaced, on a separate page(s).

Policy. Accepted manuscripts are subject to quality review concerning relevance of the article to the theme of the issue, acceptability of illustrations, reference accuracy and completeness.

The author is responsible for all statements made in the work, including changes made by the copy editor unless the author changes the latter at the time the manuscript is received for review and publication approval. Manuscripts are received with the understanding that they are not being considered for publication in another journal. Although rejected manuscripts will usually be returned to the author, *Directions* is not responsible for loss. Accepted manuscripts become the property of *Directions*, and may not be published elsewhere without written permission.

Foreword

This issue of *Directions* has two basic purposes. First, the editors hope that these articles on assessment of hearing impaired individuals will contribute to the improvement of assessment practices generally. Whether the assessment is audiological evaluation, classroom assessment, or a full psychological evaluation, the articles presented here can provide insight and, just possibly, some practical guidance for professionals in the field of deafness who are looking for new ideas.

Our second purpose is to induce you to give this journal away to a colleague in regular education. As your reading will show, many of the articles presented here have the regular class teacher in mind. Despite the passage of P.L. 94-142 half a decade ago, the link between the regular class teacher and the special education teacher remains one of the weakest links in all of education. The success of a "free appropriate public education," and placement via non-discriminatory assessment, require an unbreakable bond between the special educator and regular class teachers. The traditional suspicion and (if we are honest) sometime hostility between regular and special education may be one of the biggest unfaced problems in the implementation of P.L. 94-142. If hearing impaired students are to be properly and fully served, the cooperation of all professionals is essential, and a significant first step in achieving that cooperation is mutual understanding of one another's purposes and perspectives. The regular educator who understands how he or she can provide assessment data with high diagnostic value, and the special educator who can offer the regular class teacher sound strategies for teach-

ing, both need to learn how to talk with each other better. The authors and editors of this issue of *Directions* have that conversation in mind when we urge you to "pass it along."

Announcing . . .

The Gallaudet College Press is now completing its first year of publication and is continuing to accept manuscripts for consideration.

The Press pursues the goal of serving the deaf community through the dissemination of scholarly books, curriculum materials, literary and artistic works, and service publications.

Authors are notified of acceptance of their manuscripts after review and evaluation by the Editorial Board, comprised of individuals from throughout the country who are knowledgeable and concerned about issues of deafness.

If you have a manuscript of significance to those in the field of deafness and currently seek a publisher, please call or write for specifications:

Director
Gallaudet College Press
Kendall Green
Washington, D.C. 20002

Perhaps together we can make a substantial contribution to the deaf community.

Introduction

Where the individual hearing impaired youngster incessantly meets the education system is in the classroom. Whether the meeting of that special youngster and the teacher in that room results in benefit for the students' intellectual, social, and vocational growth depends to some relatively unknown extent upon the various methods used to assess his or her status and progress. Obviously, many other factors in the proximal and distal environment can and do effect the outcome of schooling, but what the teacher *knows* about the student and can use effectively is critical to the students' continued educational development. This vital information can come from personal observation, teacher-made tests both formal and informal, standardized tests, and the plethora of assessment information which accumulates in the records of every hearing impaired child. To maintain appropriate perspective, however, it should be remembered that a test is only *one* event and should be viewed within the context of the myriad of events yielding knowledge and experience needed to shape and direct the education enterprise.

This issue then is about assessment, i.e., assessment in its broadest sense, of the hearing impaired child. It includes articles written by various well-known authorities on issues both directly and indirectly related to the status and progress of that child. These articles range from quite detailed information on achievement testing in the classroom all the way to a hospital examining laboratory where medical evaluation might be conducted and described.

In order to convey a pretaste of the topical flavors contained in this issue, a modest content-analysis was done. Simply put, content analysis is a way of finding all the separate ideas

in a selection of prose material whether they be responses to open-ended questions or connected prose. The first step is to read the material carefully and list all of the distinct ideas as they occur. As references to a same or very similar idea recur, frequencies related to the original listed idea are accumulated. This is called "Cat-Scaling" or categorizing and ordering in frequency of occurrence. This procedure for the 11 pieces in this issue yielded 70 categories or separate ideas. Next, a careful reading of the categories lead to a possibility for combinations of related ideas. For example, the most frequently occurring idea in this issue was identified as *The Kinds of Tests or Testing* which includes ideas on: types of items such as *recall or recognition*; informal assessment; criterion referenced tests; and the modality of testing. This cluster of ideas accumulated 20 tallies and therefore underscores the importance of this aspect of the assessment topic. The *design* of an evaluation procedure should be the best possible for the fair assessment of a hearing impaired child. The provisions for *using* the test or the test administration proce-

dures accumulated nine tallies. This aspect is also crucial to fair testing. The third part of the assessment triad, that of *interpreting* the results accumulated only four tallies. This crucial aspect was not intentionally neglected, in our opinion, but the low frequency of occurrence only reflects how far we need to go in obtaining a complete picture of assessment and hearing impairment.

The cat-scaling results for the most prominent ideas are shown in Table 1.

Practical ideas like test-taking skills and training, testing for placement, and feedback to teachers occurred very infrequently. These kinds of ideas represent issues to which a lot more attention needs to be paid. The one issue which is on the horizon for handicapped children is that of competency testing. Certification of the acquisition of basic competencies for students in regular schools has already been installed in some states, and educators of hearing impaired youth should be aware of the necessity for dealing with this movement in the near future. We regret not having an article on this topic in this issue.

Table 1
Cat-Scaling Results

Frequency	Categories
20	The Kinds of Tests and Testing (a cluster)
13	Achievement Testing (a cluster)
12	The Language of Testing (a cluster)
9	Performance & Diagnostic Tests
7	Communication Skills Testing
6	Suggestions for the Tester
6	Informal Assessment (a cluster)
5	Test Administration
5	Validity
5	Intelligence Testing
5	Reading & Writing Skills Testing

Classroom Assessment Techniques for Hearing Impaired Students: A Literature Review

Barbara G. McKee
Michael A. Hausknecht

Few teachers and researchers can escape the problem of assessing student learning whether it be a unit of instructional material just used with students or a question of the efficacy of a particular instructional technique. Sooner or later one is forced to face the question: Did the students learn what was intended for them to learn? For a variety of reasons, many educators turn to some sort of paper and pencil procedure to answer this question.

This article is directed primarily to those who are interested in the use and development of classroom achievement tests for high school and college hearing impaired students. The current state of knowledge about item syntax, vocabulary, and format, as well as additional related topics is reviewed. The emphasis is on classroom tests where the content in question is other than language skills. Often the content under question—e.g., mathematics, science, machine use and safety—is generally unrelated to language skills, and the tests should measure the students' knowledge of the content presented rather than their language skills.

Although many of the findings reported in this paper are applicable to student surveys, rating scales, and opinion measures, for the most part emphasis is placed on achievement tests—i.e., those instruments designed for a right or wrong answer.

The following areas are covered in this article:

- (1) Syntactical structure and vocabulary of test items; (2) Item format; (3) Modes of delivery of test instruction and test items; (4) Test-wiseness and propensity for risk taking; (5) Nonpaper and pencil assessment techniques.

The information presented under the above headings is meant to repre-

sent current knowledge on the given topics. Studies cited, however, are meant to be representative rather than inclusive.

Syntactical Structure and Vocabulary Level

Research in this area concentrates on describing the written and spoken language "production" of hearing impaired students as opposed to a description of the syntactical structures and vocabulary that one would expect could be comprehended accurately by them.

Cooper and Rosenstein (1966) summarized 50 years of research on the language of the hearing impaired by dividing available studies into *free* and *controlled* methods of data collection. The free method characterized those studies that examined students' production of the English language, most often in a written form. They stated that the studies found deaf children wrote significantly shorter sentences, with less diversity of vocabulary and simpler sentence structure than hearing children of the same age.

Most traditional studies of the controlled type examined the performance of deaf children on standardized achievement tests. Cooper and Rosenstein (1966) noted that most studies found that the reading comprehension of deaf 18-year-olds and deaf adults rarely rose above those of the average fourth-grade hearing child.

One of the largest studies to date (DiFrancesca, 1971) has confirmed the findings of Cooper and Rosenstein (1966). Approximately 17,000 deaf students enrolled in primary and secondary programs for the hearing impaired were given the

Stanford Achievement Tests in the spring of 1971. DiFrancesca (1971) states that "Reading comprehension turns out to be the most difficult academic area of all those covered by the Stanford tests" (p. 9).

"The 1960's saw the beginning of a new series of analyses that was more concerned with intra-sentence and intra-clause phenomena and with explicating the *rules* that children use to produce sentences at various developmental levels" (Quigley, Wilbur, Power, & Montanelli, 1976, p. 25). A major project in this area was conducted by Quigley and his associates during the past several years. They tested 450 prelingually deaf students on most of the major aspects of English syntax using his *Test of Syntactic Ability (Experimental Edition)*. The particular deviations assessed were developed from earlier studies analyzing the written language of deaf children.

Quigley et al. (1976) found the most difficult structures for deaf students to comprehend accurately were pronominalization, the verb system, complementation, and relativization. The students had the least difficulty with negation, conjunction, and question formation (p. 184). Quigley et al. (1976) also pointed out that the gap that exists between the standard English of common classroom text books and the deaf children's language is very large and represents a significant impediment to their learning even the most basic content from standard texts.

Other studies (Cooper, 1967; Marshall & Quigley, 1970; O'Neill, 1973; Power, 1971; Schmitt, 1968; Taylor, 1969) have used a conceptual base akin to that used by Quigley et al. (1976) and have found results similar to those they reported. However, some researchers have taken a differ-

"... many deaf individuals have internalized grammatical systems which consist of linear linkages of lexical items ..."

ent path. Bochner (Note 1) has criticized the use of conventional grammatical models of English to characterize deviant language in deaf children and adults. He states:

In sum, describing the general principles which organize the grammar and underlie the aberrant linguistic behavior in a prelingually hearing-impaired population (or other population) entails describing the grammatical systems internalized by members of the group on their own terms and not on the basis of error typologies and/or grammatical and acquisition models of standard English. (p. 5)

Bochner hypothesized that many deaf individuals have internalized grammatical systems which consist of linear linkages of lexical items and do not specify sub- or superordination. Bochner (Note 1) dubbed this theory "The Linear Order Hypothesis," and has completed a study testing it. He summarized his findings by saying:

The data indicate that many deaf individuals do not have an adequate knowledge of grammatical relationships which may occur among clauses, consequently, the individuals tend to arrange clause in a linear-sequential fashion without specifying relationships of subordination. (p. 30)

He qualified this finding by saying his study indicated that approximately $\frac{1}{3}$ of his subjects could be described by this hypothesis, slightly less than $\frac{1}{3}$ could not, and slightly more than $\frac{1}{3}$ of the study group did not fit the parameters.

Several studies have examined the vocabulary level of deaf students separately from sentence syntax. In their extensive review, Cooper and Rosenstein (1966) stated that the

vocabulary levels of deaf 18-year-olds was judged to be about the fourth-grade level. Gentile and DiFrancesca (1972) found similar results in the nationwide testing program conducted by the Office of Demographic Studies at Gallaudet College.

Other studies have also found that hearing impaired people are restricted in terms of flexibility and breadth of vocabulary usage (Rosenstein & MacGinitie, 1969; Silverman-Dresner & Guilfoyle, 1972). Walter (1977, 1978) developed a technique to measure the vocabulary comprehension of deaf students based on the idea that their vocabulary knowledge can be directly related to the frequency of occurrence in print of any particular word. He used the *American Heritage Word Frequency Book* which lists, in order of frequency of occurrence, over 25,000 words. Walter (1977, 1978) found that for 14-year-old deaf students to read correctly 66% of the words given them, the words must fall within the 2,000 words which appear most frequently in the *Heritage List*.

Use of this technique with college-age students indicated that in order to maintain a mean of 66% correct interpretation, the words must fall within the 10,000 level in frequency.

In summary, this section offers some suggestions to the classroom teacher attempting to construct paper and pencil tests of subject matter not related to reading and writing skills:

1. Use a vocabulary level appropriate for pupils:
 - (a) Junior high age—2,000 on the American Heritage list.
 - (b) College age—10,000 on the American Heritage List.
2. Avoid the following grammatical structures:

- (a) Compound and complex sentences in test items and directions.
- (b) Wh-questions and complex negation formation.

Item Format

Researchers interested in the problems of assessing the knowledge of deaf students have only recently shown interest in the problems concerned with the item format used in tests with deaf students. Trybus and Buchanan (1973), while examining patterns of achievement in a nationwide administration of the Stanford Achievement Test, noted two special problems in the course of their analysis.

Position of the correct response

Traditionally, in multiple-choice tests the position of the correct answer is randomized among the alternatives. It is expected that the performance of the groups will not be affected by the positioning of the correct response. This was not found in Trybus and Buchanan's (1973) analysis. While the specific position showing the best or poorest performance among hearing impaired students varied from subtest to subtest on the Stanford Achievement Test, there seemed to be a definite difference due to positioning. Trybus and Buchanan (1973) stated, "Some experimental study and the evaluation of various kinds of test-taking training will be needed to clarify the issues raised by these findings" (p. 18).

Negative responses

The second special problem noted during the Trybus and Buchanan analysis is that of the students' responses to items in which the correct

"... various methods of presenting instructions for a test ... have not been extensively examined in the literature ..."

answer is a negative answer. The forms of the negative responses were of the following types:

"N"—Neither alternative is correct.

"NP" No punctuation needed.

"NG"—Correct answer is not given.

Hearing impaired students tended to do substantially less well than their hearing peers on items where the correct answer was one of the above "negative" responses.

Matching vs. multiple-choice test format

Walter (1976) administered two forms of a vocabulary test to two groups of entering National Technical Institute for the Deaf (NTID) students during the summer of 1976. One form of the test used an 8-sentence, 10-word matching format. The other form employed the same words and sentences in a multiple-choice format with five options.

Student responses on the two formats were compared. Walters (1976) found that the mean scores for the matching format and the multiple-choice format did not differ significantly. Hence, the use of either format did not appear to influence the students' total scores.

Recall vs. recognition items

A study by Osguthorpe, Long, & Ellsworth (Note 3) measured the effects of reviewing class notes on four learning outcomes: recognition, recall, concept acquisition, and problem solving, for a group of 53 deaf and 33 hearing college-age students. An unexpected result that surfaced in the study was the performance of the deaf students on the recall and recognition items.

The individual means show that hearing students nearly doubled

their performance from recall ($m=5.8$) to recognition ($m=10.1$) while deaf students increased only slightly, recall ($m=4.48$), recognition ($m=6.05$). Further analysis showed that approximately half (47%) of the delayed review deaf students' recognition scores were equal to or actually lower than their recall scores. . . . In other words having the response set available was highly facilitative for hearing students, but not for deaf students. Many deaf students were actually confused by having the correct answer on the form. This finding has been replicated in a subsequent experiment (Osguthorpe et al., 1977, pp. 11-12).

Item bias

In a recent article, Rudner (1978) applied a technique suggested by Angoff (1972) to determine which items contained in 12 subtests of the Stanford Achievement Tests: Intermediate I battery were biased for or against deaf students.

Rudner (1978) identified 48 items (of 419) that were biased in favor of one population or the other. Of these 48 biased items, 22 appeared to be biased in favor of hearing impaired examinees and 26 in favor of hearing examinees. Rudner (1978) stated that since the number of biased items is low and since the biased items are fairly uniformly distributed between population groups, one may suspect that there is little bias in total scores. However, he pointed out two reasons this may not be the case:

1. The limitations of the procedure may lead to spurious conclusions. Indices of item difficulty are obtained for each item under investigation for each group. These difficulty levels are plotted on a 2-axis graph.

Items deviating significantly from the main axis of the plot are considered to be biased. Item difficulty may not be the only characteristic that should be considered in the determination of item bias.

2. In this study, the hearing impaired subjects were, on the average, five or six years older than the hearing subjects. Students were assigned to a particular battery based on reading levels. Because of their higher mean age (and hence greater repertoire of experience), the hearing impaired students might do well even on items biased against them.

Rudner (1978) found that the following linguistic structures appear to be misleading for hearing impaired students:

1. Conditionals (if, when);
2. Comparatives (greater than, the most);
3. Negatives (not, without, answer not given);
4. Inferentials (should, could, because, since);
5. Low information pronouns (if, something);
6. Lengthy passage.

Methods of Delivery

The strength and weakness of various methods of presenting instructions for a test as well as the test items themselves have not been extensively examined in the literature dealing with assessment of the hearing impaired. Most studies in this area have concentrated on examination of different modes of presenting learning materials; usually the criterion measure is some sort of paper and pencil test. Unfortunately, the

"Risk taking . . . is the tendency to respond to a test item when the correct answer is not known . . ."

psychometric properties of the criterion test are not always examined.

Stuckless (Note 4), in a paper discussing the application of technology to deaf education, reviewed several studies that investigated different visual modes of presenting verbal information to deaf persons (Gates, 1971; Norwood, 1976; Propp, 1972; Stuckless & Pollard, 1977). Stuckless (Note 4) concluded that "The results of both basic and applied research . . . furnish conditional support for the use of print for receptive communication by deaf persons" (p. 11).

Higgins (1973) compared performance of three groups of Gallaudet students on a multiple-choice test after they had viewed two brief videotapes: one group used total fingerspelling (The Rochester Method); one group used American Sign Language (ASL); and the third used Siglish, a form of manually coded English. Siglish was found to lead to significantly better performance on the criterion measure than either ASL or fingerspelling. It was pointed out that the criterion measure was written in standard English (thus Siglish more closely approximated the criterion measure), and Gallaudet students are generally accustomed to some form of manually coded English through their classroom work.

Hoemann (1972) wanted to test the accuracy of interpreted communication in a group situation. At a workshop for interpreters for the deaf sponsored by the Ohio Chapter of Interpreters for the Deaf, a simple 10-item test was both read aloud and signed to the participants. All participants were adults; the deaf participants were invited consultants who were respected by both the deaf

community and the Interpreter Organization (an average or above-average intelligence for these subjects can be assumed). The questions were such that any adult with normal intelligence would have little difficulty in answering them correctly, e.g., in what month does Christmas fall (one nontrivial question was included to lend credibility to the statement that this was a test). The mean score of the 16 hearing subjects was 9.1 ($n=16$), the score of the hearing impaired participants was 5.6 ($n=9$). Hoemann's (1972) study points out the problems of receiving manual communication in a group setting even when the receivers are motivated adults.

"Test-Wiseness" and Risk Taking

Two aspects of testing receiving quite a bit of attention in the literature concerned with factors affecting or test scores for the hearing population have been those of "test-wiseness" and risk-taking behaviors. This area has received almost no coverage in the literature concerned with assessment of deaf students. However, the study of Koehler, Kelly, and Gonter (Note 2) examined these phenomena in a hearing impaired population. Test-wiseness (TW) is defined in the Koehler et al. (Note 2) study as "a subject's capacity to utilize the characteristics and formats of the test and/or test taking situation to achieve a high score."

Risk taking (RT), on the other hand, is the tendency to respond to a test item when the correct answer is not known. According to Koehler et al. (Note 2) there are data that indicate low-risk takers tend to be penalized on their test scores com-

pared to their more adventuresome associates.

Koehler et al. (Note 2) found that deaf students were significantly lower in TW ability than their hearing counterparts and that the degree or amount of TW did not increase with grade level or reading level as it did with hearing students. However, they did consistently receive scores above the chance level, indicating that some use of test-wiseness was being applied.

Differences in the propensity for risk taking was much smaller between deaf and hearing subjects. Although ninth-grade hearing subjects were substantially higher-risk takers, the difference became less in the 10th and 11th grades and disappeared by the 12th grade.

Nonpaper and Pencil Assessment Techniques

One of the least examined (or at least unpublished) topics in the general area of assessing classroom learning of deaf students has been that of nonpaper and pencil assessment devices. Certainly, teachers have used various types of observation techniques to determine when their students have mastered certain skills, but such techniques have rarely been systematically examined and traditional achievement tests are even more rarely compared to other methods of assessing subject mastery.

Vernon and Brown (1964) strongly recommend that all I.Q. tests administered to deaf children be non-verbal performance instruments. They have backed up their recommendation with examples of children severely misdiagnosed on the basis of verbal I.Q. tests. A review of 50

"Why did the students choose a particular response or give a particular answer?"

years of research on intelligence testing of deaf persons (Vernon, 1968) reinforces the fact that when deaf children are tested by appropriate methods and have no additional handicapping condition the deaf and hard-of-hearing function at approximately the same I.Q. level on *performance* intelligence tests as do the hearing.

The NTID Visual Communication department has developed an extensive and relatively objective student assessment plan for their Printing Production Program that does not rely on paper and pencil assessment instruments. The faculty developed specific skill objectives and required that the students demonstrate mastery of the skills in a laboratory situation. Such a technique is similar to the competency-based curriculum approach now receiving so much attention in higher education. One aspect of occupational education that might be considered different from the liberal arts is the need for the student to reach a performance level which approximates 100%. Hence, the students are simply rated on whether or not they can successfully complete the task; the only variable is the time needed to reach that level of performance. Such time data are now routinely collected in the Printing Program.

Preliminary results of the time data indicate a possible bimodal distribution. "There is an implication that students who do not complete a skill or skills within the normal time distribution will never complete a course unless something is done to change the students' learning success" (Visual Communications Department, Note 5, p. 30). Most certainly this is an area that deserves additional examination.

Summary and Suggestions for Test Developers

Suggestions arising from the literature reviewed include:

1. Confine tests to a vocabulary level appropriate for the pupils; for high school and college-age students, this would be somewhere between the 5,000 and 10,000 word level on the *American Heritage Word List*.
2. Whenever possible, avoid compound and complex sentences in test items and test directions. Minimize the number of wh-questions and complex negation formations.
3. Many students will be interpreting phrases in a linear fashion. Take this into account when developing test items.
4. Minimize items in which the correct answer is a negative answer.
5. Be aware that a multiple-choice or matching format (recognition) item may not be easier for students than a fill-in-the-blank or a short-answer item requiring a one-word response.
6. Caution students to read the entire item (all options) carefully before responding.
7. Always include complete written instructions. Use manual communications to reinforce and elaborate on the written instructions.
8. If possible, use visual aids to avoid lengthy narrative descriptions in both test directions and test items.

The research conducted thus far concerning the test wiseness and risk-taking characteristics of hearing impaired students is so minimal that no conclusions or suggestions can be

offered. The final topic covered in this paper is nontraditional methods of assessing the learning of deaf students. Again, so little has been documented that no conclusions can be drawn or recommendations offered. Clearly, this is one area that urgently needs further study.

In summary, the authors would like to make a recommendation to test developers. Conduct a pilot test for any classroom assessment technique that will be used repeatedly or that will carry above-average importance. After the pilot test, the students' responses should be carefully examined and the test or assessment technique should be discussed with them. Questions concerning the test might include: How did the students interpret a given question? Why did the students choose a particular response or give a particular answer?

Suggestions for Future Research

The research reviewed for this paper raised more questions in the minds of the reviewers than it answered. In general, any additional studies in the area covered by this review should be of great significance to those concerned with the educational assessment of hearing impaired individuals. The authors offer the following comments:

1. Bochner (Note 1) as well as others should be strongly encouraged to continue the investigation of the Linear Hypothesis Theory.
2. More study is needed on the vocabulary comprehension skills of hearing impaired students. For example, is the *American Heritage Word List* the most feasible instrument to use or

should a modification of such a list be developed for the use of those designing instructional materials and assessment techniques for use with the hearing impaired? Deaf students' comprehension of second and third meanings of words need further exploration.

3. More studies which examine the effect of the positioning of the correct response in multiple-choice questions need to be conducted.
4. Further work should be undertaken examining the item formats in objective tests.
5. The test-wiseness ability of deaf students should be studied further.
6. Congruence between information delivery methods and methods of assessing hearing impaired students should be studied. Several studies have examined different modes of presenting information, but the criterion measure was almost always a paper and pencil test in standard English.
7. More methods employing the demonstration of desired skills need to be developed and examined. For example, the NTID Printing Program makes the assumption that a student either can or can not perform a given task. Is this achievement information a truly dichotomous variable?

Reference Notes

1. Bochner, J. *Error, anomaly, and variation in the English of deaf individuals*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, 1978.

2. Koehler, R. A., Kelly, R. R., & Gonter, M. A. *Test-wiseness and risk taking of deaf and normal hearing subjects*. Paper presented at AERA Convention, New York, 1977.
3. Osguthorpe, R. T., Long, G. L., & Ellsworth, R. G. *The effects of reviewing class notes for deaf and hearing students*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, undated.
4. Stuckless, E. R. *Technology and verbal processing*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, undated.
5. Visual Communications Department. *A plan of work to implement individualization of instruction in printing production*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, 1976.

References

- Angoff, W. *A technique for the investigation of cultural differences*. Paper presented at the annual meeting of the American Psychological Association, Honolulu, May 1972.
- Cooper, R. The ability of deaf and hearing children to apply morphological rules. *Journal of Speech and Hearing Research*, 1967, 10, 77-78.
- Cooper, R. L., & Rosenstein, J. Language acquisition of deaf children. *Volta Review*, 1966, 68, 58-67.
- DiFrancesca, S. Academic achievement test results of a national testing program for hearing impaired students—United States: Spring,

1971. *Annual Survey of Hearing Impaired Children and Youth*, D, (9), 1-41.

Gates, R. The reception of verbal information by deaf students through a television medium: A comparison of speechreading, manual communication, and reading. *Proceedings of the Convention of American Instructors of the Deaf*, 1971, pp. 513-522.

Gentile, A., & DiFrancesca, S. *Item analysis of academic achievement tests hearing impaired students United States: Spring 1969*. Gallaudet College, Office of Demographic Studies, Series D, Number 2, Washington, D.C., 1972, 1-34.

Higgins, E. An analysis of the comprehensibility of three communication methods used with hearing impaired students. *American Annals of the Deaf*, 1973, 118, 46-49.

Hoemann, H. W. Communication accuracy in a sign-language interpretation of a group test. *Journal of Rehabilitation of the Deaf*, 1972, 5, 40-43.

Marshall, W., & Quigley, S. P. *Quantitative and qualitative analysis of syntactic structure in the written language of deaf students*. Urbana, Ill.: Institute for Research on Exceptional Children, 1970.

Norwood, M. *Comparison of an interpreted and a captioned newscast*. Unpublished doctoral dissertation, University of Maryland, 1976.

O'Neill, M. A. *The receptive language competence of deaf children in the use of the base structure rules of transformational-generative grammar*. Unpublished doctoral dissertation, University of Pittsburgh, 1973.

Power, D. H. *Deaf children's acquisition*

- tion of the passive voice. Unpublished doctoral dissertation, University of Illinois, 1971.
- Propp, G. *An experimental study on the encoding of verbal information for visual transmission to the hearing impaired learner*. Unpublished doctoral dissertation, University of Nebraska, 1972.
- Quigley, S., Wilbur, R., Power, D., Montanelli, D., & Steinkamp, M. *Syntactic structures in the language of deaf children*. Institute for Child Behavior and Development, University of Illinois at Urbana-Champaign, 1976.
- Rosenstein, J., & McGinitie, W. (Eds.). *Verbal behavior of the deaf child: Studies of word meanings and associations*. New York: Teachers College Press, 1969.
- Rudner, L. M. Using standard tests with the hearing impaired: The problem of item bias. *Volta Review*, 1978, 80, 31-40.
- Schmitt, P. *Deaf children's comprehension and production of sentence transformations and verb tenses*. Unpublished doctoral dissertation, University of Illinois, 1968.
- Silverman-Dresner, T., & Guilfoyle, G. *Vocabulary norms for deaf children*. New York: 1972.
- Stuckless, E. R. & Pollard, G. Processing of fingerspelling and print by deaf students. *American Annals of the Deaf*, 1977, 122, 475-479.
- Taylor, L. *A language analysis of the writing of deaf children*. Unpublished doctoral dissertation, State University of Florida, 1969.
- Trybus, R. J., & Buchanan, C. Patterns of achievement test performance. *Studies in achievement testing, hearing impaired students United States: Spring, 1971*. Gallaudet College, Office of Demographic Studies, Series D, Number 11, Washington, D.C., 1973, 3-44.
- Vernon, M. Fifty years of research on the intelligence of deaf and hard-of-hearing children: A review of literature and discussion of implications. *Journal of Rehabilitation of the Deaf*, January 1968, pp. 1-12.
- Vernon, J., & Brown, D. W. A guide to psychological tests and testing procedures in the evaluation of deaf and hard-of-hearing children. *Journal of Speech and Hearing Disorders*, 1964, 29, 414-423.
- Walter, G. *A test to measure vocabulary knowledge of deaf students*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, 1977.
- Walter G. *Lexical abilities of hearing and hearing impaired children*. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, 1978.

Informal Assessment of Hearing Impaired Students in the Classroom

Barry R. Culhane
Ron Hein

Teachers of hearing impaired students periodically evaluate their students with formal tests, but they are constantly evaluating their students on an informal basis during classroom interactions. Rather than focusing on the traditional, formal, structured tools of assessment (e.g., classroom tests and standardized tests), this article focuses on that informal assessment and communication process and its relationship to instructional goals.

There are three main purposes of informal assessment: (a) to provide students with a guide to teacher expectations of what will be required in the classroom, (b) to provide the teacher with a measure of students' learning relative to goals set, and (c) to stimulate active learning processes in the students. When one considers the activities which occur in the classroom, it becomes apparent that informing students of the classroom goals and expectations is often a subtle process. As information is provided to students and probing questions are asked, a teacher is also providing students with cues about what is expected from them in the classroom.

If a teaching style of lecturing to students and asking questions which lead to only one correct answer (e.g., Q: "Who discovered America?" A: "Columbus") is adopted, students learn that their task is to memorize factual information. Students become accustomed to using only closed, convergent thinking processes. If a teacher employs a style combining lecture with discussion groups and asks questions with several potentially "correct" answers (e.g., Q: "Who was the most significant early explorer of America and why?" A: Many different possibilities), students may learn that their task is to

analyze, synthesize, and evaluate information by using divergent, open-ended thinking processes. Obviously, the nature of the probing will depend upon the instructional goals of the lesson. Both convergent and divergent approaches are valid for different goals. The author's intent is to encourage the use of informal assessment to stimulate active learning, not to urge the use of one style of questioning to the exclusion of another.

Another primary function of informal assessment is to provide the teacher with the feedback needed to assess students' progress. In reality, both functions (i.e., stimulating active learning and assessing students' progress) are part of one interactive process of communicating effectively in the classroom. Too often, teachers have been surprised by poor formal test results because they incorrectly interpreted reassuring nods (and other demonstrations of student approval) to mean that students were fully comprehending the instruction. Student responses and questioning behaviors give some of the best cues on the scope and depth of learning occurring in classes. When those behaviors are lacking, it should be a cue that something is amiss.

The third primary function of informal assessment is to promote active learning among students. Active learning is a process which can occur when a teacher is challenging his or her other students to seek out information and to locate and manipulate ideas and relationships for themselves. Active learning is contrasted with passive learning where the expectation is that both questions and answers are given for absorption by students. Active learning involves the idea mentioned earlier of asking divergent questions rather than ex-

clusively convergent (or "one answer") questions. It also involves the use of higher level cognitive questions such as those discussed by Bloom, Engelhart, Furst, Hill, and Krathwohl (1956). In stimulating active learning, the instructor requires the participation of students in classroom discussions and the production of ideas by these students. If active learning is to occur, you cannot expect students to simply recite what you have told them. They must produce something which is new to themselves. This is not a simple process; it requires work from both teachers and students.

One useful way to make informal assessment work is to analyze how it occurs in the classroom. To assist teachers in this process, the authors are sharing a system for categorizing questions and a description of teacher and student activities which encourage various kinds of learning. In addition, a summary of strategies to facilitate higher order thinking processes is presented. The assumption is that factual information needs to be learned throughout life, especially at the initial stages of new learning. However, all students (including hearing impaired students) need to develop an ability to evaluate new information from multiple perspectives and learn to value their own ideas and opinions. It is suggested that informal assessment in the classroom can play a major role in meeting that goal.

Stimulating Active Learning through Questions and Observations

There are two primary ways for both teachers and students to promote active learning in the class-

"Sometimes the apparent daydreamer is really thinking about the subject matter."

room. The first is through observational activities, and the second is through question asking and answering activities:

Teacher's observational activities

A teacher's observational activities are a key component of informal assessment. The goal of these observational activities is to intentionally search for student feedback about where students are in relation to the instructional goals.

It is believed that certain behaviors are common and have specific implications for the instructional process. Among the more common student behaviors signalling a need for probing questions are: looks of puzzlement, lack of response, disruptive behavior, frequent requests for information from other students, and daydreaming. What do these imply? They imply that the student has a problem which is not being adequately expressed. Assuming that the problem is not one of pure boredom or lack of motivation, it is a mutual problem of ineffective classroom communication. How can a teacher locate and resolve the problem? Possibly through informal assessment and one key teacher activity, question asking. After observing a potential problem, a teacher should check his or her basic assumptions about why the behavior is occurring. Sometimes the apparent daydreamer is really thinking about the subject matter. For teachers seeking explicit information on observational techniques, the collection of articles edited by Weinberg and Wood (1975) contains a good basic set of concepts and procedures.

Questioning activities for teachers

What are different types of questions that can be asked in a class-

room? The authors recommend using a questions classification scheme developed by Blosser (1973) for classifying questions typically asked by teachers and by students. Among the several classification systems developed, most widely known is that of Bloom et al. (1956) which is used for classifying cognitive domain activities. Other related schemes developed by Guilford (1967), Warburton (1970), and Gagne (1970) are wide-

ly used in psychology for the classification of both cognitive processes and intelligence test questions. However, the authors have found that the scheme developed by Blosser (1973) includes many of the considerations of these systems, and it additionally offers a structure for classifying all types of classroom questions. These features are readily apparent as one examines the scheme as it is represented in Figure 1.

Figure 1
The Question Category System
(Blosser, 1973, p. 54).

Question Category System for Science		
Level I	Level II	Level III
I. CLOSED QUESTIONS (limited number of acceptable responses)	A. COGNITIVE MEMORY*	1. RECALL: includes repeat, duplicate, memorized definitions 2. IDENTIFY or NAME or OBSERVE
	B. CONVERGENT THINKING*	1. ASSOCIATE and/or DISCRIMINATE; CLASSIFY 2. REFORMULATE 3. APPLY: previously acquired information to solution of new and/or different problem 4. SYNTHESIZE 5. CLOSED PREDICTION: limitations imposed by conditions or evidence 6. MAKE "CRITICAL" JUDGMENT: using standards commonly known by class

“... the goal of questions should be twofold: evaluating the learner’s progress and stimulating the student to be an active learner.”

Figure 1 (continued)

Level I	Level II	Level III
II. OPEN QUESTIONS (greater number of acceptable responses)	C. DIVERGENT THINKING*	1. GIVE OPINION 2. OPEN PREDICTION: data insufficient to limit response 3. INFER or IMPLY
	D. EVALUATIVE THINKING*	1. JUSTIFY: behavior, plan of action, position taken 2. DESIGN: new method(s), formulate hypotheses, conclusion(s) 3. JUDGE A: matters of value, linked with affective behaviors 4. JUDGE B: linked with cognitive behaviors
III. MANAGERIAL Teacher uses to facilitate classroom operations, discussion		
IV. RHETORICAL Teacher uses to reinforce a point; does not expect (or want) a response		
*1. <i>Cognitive-Memory</i> : evidence understood to be directly available (textbook, previous lesson or discussion, film, filmstrip, chart, experiment, field trip, etc.) 2. <i>Convergent Thinking</i> : evidence directly available but not in the form called for by question 3. <i>Divergent Thinking</i> : evidence for response not directly available 4. <i>Evaluative Thinking</i> : evidence may or may not be directly available; criteria for responding available, directly or indirectly. Implication that student may be called upon to provide a defense for his response.		

The Blosser (1973) scheme highlights the open or divergent and the closed or convergent questions mentioned, along with two other types commonly found in most classrooms—managerial and rhetorical questions. This scheme should be helpful in analyzing the questions which currently are asked during classroom interactions.

Some specific examples of questioning activities containing both the stimulation and the evaluation components useful for the instructional process are provided in this article. The key point to remember is that the goal of questions should be twofold: evaluating the learner’s progress and stimulating the student to be an active learner. It is important to ask questions appropriate to the instructional goal and not reduce the level of the question to get a response. Patience in the face of silence when asking for a response communicates to students that the teacher wants a response which requires thinking. If teachers start with a complex question (“What were the causes of World War II?”) and reduce the question to a closed question or no question (“Did Germany invade France?”) they may be more likely to get an answer, but they have also told the students they will settle for less than active thinking. Table 1 contains various questions, classified as either open or closed, according to the Blosser (1973) scheme. While it is not felt that open questions should be used to the exclusion of closed questions, it is believed that open questions lead to more active learning. Elsewhere, Long, Hein, and Coggiola (1978) and Hein (1979b) have presented data on the value of active learning processes with hearing impaired students.

"What can be examined by informal assessment?"

Table 1
Examples of Active, Higher-Level Questions,
Contrasted to Lower-Level Questions

Questions	Category	Goal
(1) Open: When is a computer more valuable than a book for storing information? Closed: What are two ways to store information?	Evaluation/ justification Cognitive memory/recall	Combining ideas Recalling facts
(2) Open: If you were given the height of a container, could you calculate its volume? Closed: What is the formula, base times height times width used for?	Evaluation/ judgment Cognitive memory/recall	Linking cognitive behaviors Recalling facts
(3) Open: Is a movie or a trip to the zoo more fun? Closed: What did you see at the movies? What did you see at the zoo?	Evaluation/ judgment Cognitive memory/name	Giving values Giving names and details
(4) Open: How could you teach Jim the path from your house to school? Closed: How do you walk from your house to school?	Evaluation/ design Cognitive memory/recall	Designing a strategy Recalling details

What can be examined by informal assessment?

Types of behaviors to look for during informally observing of students, and a question classification system for questions asked by both teacher and students have been discussed. Next, some of the things which can be informally observed and questioned in a classroom will be listed. The following can be examined through informal assessment:

1. Screening for placement within programs, reading/math/etc.;
2. Evaluating student's current

status relative to teacher's goals (academic, personal/social, communication, physical);

3. Determining the effectiveness of instruction;
4. Determining the student's rate of learning;
5. Determining the student's preferred mode of learning (auditory, visual, motor, tactile, sign, total communication);
6. Determining the student's learning methods and strategies;
7. Determining the optimal de-

gree of classroom control or structure;

8. Determining the utility of certain materials and methods of instruction (and testing) relative to student's modes of learning;
9. Locating social variables (seating problems, peer interactions, student and teacher interaction patterns);
10. Locating SES or cultural differences or similarities that might show the need to either modify examples, materials, or methods or even modify the goals of instruction;
11. Determining physical space accommodations which are necessary to promote an optimal learning environment (lighting, acoustics, traffic patterns, furniture placement, seating positions);
12. Discovering parental/family interactions and needs;
13. Discovering student's interaction problems/strengths in nonclassroom settings or in other classroom or school settings;
14. Determining which activities most successfully interest and motivate students; and
15. Assisting students in understanding their own feelings, attitudes, values, and opinions about key issues or events.

Reynolds and Birch (1977) provide information on traditional assessment approaches within many of these settings and on many of these topics. However, it is suggested that questioning and observation also allow informal assessment of these in most classrooms.

Student observational activities

As was suggested with teachers'

"Encouraging open inquiry by students leads to creativity in the use of a variety of cognitive skills . . ."

observational activities, students' observational activities should be focused not only on their fellow students but also on their teachers. Such observation can lead to modeling of behavior, which is an effective method of learning.

Teachers try to cue students into the important aspects of their instruction. They underline points on the board; they repeat information several times; they emphasize key concepts; they build from the simpler to more complex and interrelated concepts; they review and summarize at the end of one class and at the start of the next; they give work sheets, advance organizers, pep talks, and quizzes. Students who observe these behaviors should be cued into the fact that key concepts and relationships are being presented and that they are expected to do something with them. Students should make certain that they understand the concepts and relationships by asking themselves questions of the various types outlined previously in Figure 1. They should become active in the process of learning.

Students should also observe the questions and behaviors of their peers. They should try to understand why another student is confused or has asked a particular question and how the response to that question fits into the whole lesson. Students who do this are themselves active, involved learners. A teacher can help students develop this ability by communicating personal observations and encouraging students to do so themselves.

Student questioning activities

For students, part of the active learning process should focus on their personal behaviors both in the classroom and in study sessions. Hearing

impaired students can be taught to use both questioning and observational techniques to promote their learning and to aid them in clarifying classroom situations and learning. Teachers should not be the only users of the questioning and observational techniques outlined.

A question-asking training program has been developed (Hein, 1979a) which can be used to highlight the component parts of question asking. The process also encourages students to ask questions both during classes and as a learning strategy for improving their comprehension and retention. The training program sequentially presents students with (a) general background information on question asking; (b) information on types of questions and their relative strengths and weaknesses; and (c) the levels of questions, in terms of a systematic classification system. Although the program is primarily focused on question asking as a learning strategy, its modification into a question-asking procedure for use in a classroom is quite straightforward.

To modify the procedure for classroom use, students should be taught the concepts as noted previously and then be given illustrations of the questions, categories, and goals as outlined in Table 1. It must be stressed to the students that the goal of their question asking is to gain systematically more and more higher-order cognitive information from both their peers and their teacher: by becoming active, involved learners. Encouraging open inquiry by students leads to creativity in the use of a variety of cognitive skills, and opens the way for students to express their own opinions. Students need an opportunity to check their own values and attitudes with teach-

ers and peers through classroom interaction, small group discussion, or daily journals of their own writings or drawings. Games that reward question-asking behavior can help transfer good questioning skills by students to the regular classroom setting.

Summary

Basic ideas on informal assessment for both teachers and hearing impaired students have been presented. A question classification scheme for the use of teachers has been described; the scheme can be used to promote greater awareness of the purposes of asking questions in the classroom. A list of situations and variables which can be at least partially examined through informal assessment has been given. It has been suggested that observational and question asking/answering activities are valuable informal assessment functions in classroom situations since these can be used to accomplish three basic goals: (a) to provide students with a guide to teacher expectations, (b) to provide the teacher with a measure of the student learning relative to the instructional goals, and (c) to stimulate active learning processes in students. Informal assessment is a form of communication which is a powerful tool in shaping the nature of the teaching/learning process and the student's active role in that process. However, informal assessment should not be used in isolation from formal assessment; these are complementary procedures, both of which are needed for successful learning and instruction.

References

- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). *Taxonomy of educational objectives. The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay, 1956.
- Blosser, P. A. *Handbook of effective questioning techniques*. Worthington, Ohio: Education Associates, 1973.
- Gagne, R. M. *The conditions of learning* (2nd ed.). New York: Holt, Rinehart, & Winston, 1970.
- Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.
- Hein, R. D. *Question asking: A new way to learn more as you read and study* (Technical Manual). Rochester: Rochester Institute of Technology, National Technical Institute for the Deaf, 1979. (a)
- Hein, R. D. *Question asking: A technique for facilitating the comprehension and retention of instructional material*. Paper presented at the annual meeting of the American Educational Research Association, April 1979. (b)
- Long, G. L., Hein, R. D., & Coggiola, D. C. *Networking: A semantic-based learning strategy for improving prose comprehension* (Tech. Rep., Paper Series No. 26). Rochester: National Technical Institute for the Deaf, September 1978.
- Reynolds, M. C., & Birch, J. W. *Teaching exceptional children in all America's schools*. Reston, Va: Council for Exceptional Children, 1977.
- Warburton, F. W. The British intelligence scale. In W. B. Dockrell (Ed.), *On intelligence: The Toronto symposium 1969*. London: Methuen & Co., 1970.
- Weinberg, R. A., & Wood, F. H. (Eds.). *Observation of pupils and teachers in mainstream and special education settings: Alternative strategies*. Reston, Va: Council for Exceptional Children, 1975.

Educational Assessment for Placement

Robert K. Lennan

When a regular educator is asked to enroll a deaf child in a school program, it may be his or her first contact with a deaf youngster. This person may feel at a loss in trying to determine what educational placement would be in the child's best interests. In placing a handicapped child in the least restrictive appropriate environment, the critical question is "Which educational placement will most appropriately serve the unique needs of this child?"

For a child who is deaf, this question takes on even greater long-range importance because of the impact that deafness has on language development and on the whole learning process. Inappropriate educational placement can have a devastating and far-reaching effect on deaf children's future educational, social, and psychological development.

The following criteria for determining whether or not a deaf student will gain maximum profit from a part-time assignment to a regular class have been suggested by Hayes and Griffing (1967):

1. The student should be able to participate at or near the grade level of the regular class in using the receptive and expressive skills—speechreading, speech, language, reading, and writing.
2. The student's level of social and emotional maturity should be at least equal to that of the students in the regular class to which the assignment is to be made.
3. The student should give attention to the job at hand and follow directions well.
4. The student should be sufficiently independent, self-confident, and determined to function successfully in the regular class program.

5. The student's ability to learn, as indicated by the results of a standardized test, should be average or above average.
6. The student's chronological age should be within two years of the average age of the students in the regular class or classes for which part-time assignment is considered.
7. The students in the regular class should be ready to accept the deaf student as a member of the class and treat him or her with respect and consideration.
8. The teacher of the regular class should understand the problems faced by a deaf student assigned to the regular class and be prepared to help the student solve problems.
9. The enrollment of the regular class should be sufficiently limited to permit the teacher to have opportunity to provide the special help needed by deaf students.
10. Appropriate sound amplification should be available for use in regular classes.
11. The family of the deaf student should be interested in having him or her assigned to regular classes, helping with home assignments as much as it is possible and advisable to do so, and helping to solve any problems that may be encountered in adjusting to the regular class environment.
12. The deaf student should be willing and reasonably eager to accept an assignment to a regular class or to more than one such class.

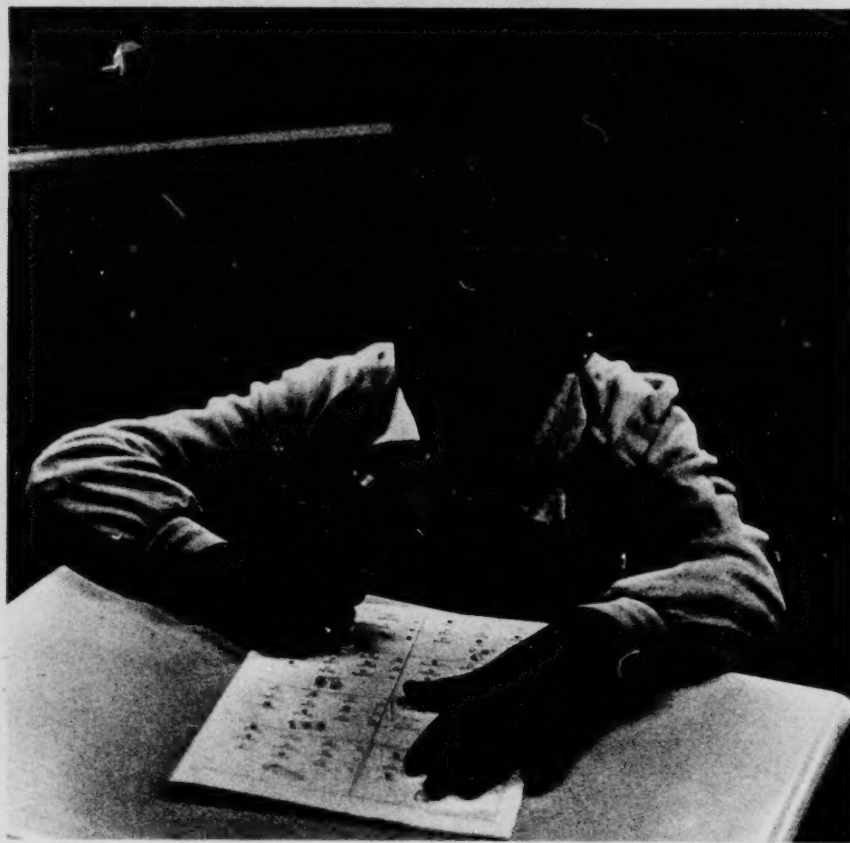
Generally, the child who has had a severe to profound hearing loss since birth or during the first two years of life will have difficulty in functioning

in a regular class placement because he or she lacks the communication skills and social development essential for sharing ideas, information, and experiences in the instructional program in regular classes (Hayes & Griffing, 1967).

How can it best be determined whether or not a deaf youngster has the communications skills necessary for successful integration into a regular classroom? Assessment activities that may be useful include the following:

1. **Hearing/Speech/Speechreading**
 - (a) Prior to assessment of the child, his or her audiogram should be reviewed with an audiologist or an otologist who should be able to provide substantial information concerning degree of hearing loss and potential ability of the child to understand speech in the classroom environment.
 - (b) Before conducting the educational assessment, it should be determined: (1) whether a hearing aid is used, (2) whether the child is wearing the aid, and (3) if it is working properly.
 - (c) During assessment of the child, the evaluator should sit facing him or her at eye level and carry on a conversation in a normal tone of voice, with normal mouth movements. Voices should not be raised and mouth movements should not be exaggerated. The assumption should not be made that because the youngster nods and smiles, he or she understands, because deaf children learn early in life to nod and smile as people talk to them and thus avoid embarrassment.

"... it can be assumed that the student is a fairly proficient speechreader."



(d) If attempts at conversation have been successful, the evaluator should continue to talk with the youngster in a normal tone of voice while covering his or her mouth with a card or piece of paper. If there is a reduction in the level of understanding by the youngster, the assumption can be made that he or she has some proficiency in speechreading, using it in combination with residual hearing to understand what is said by others. If there is no

significant reduction in understanding, then it would appear that the youngster is able to comprehend through residual hearing alone.

(e) The evaluator can make a further check on speechreading ability by talking with normal mouth movements but without voice. If the youngster is able to understand what is said without too much difficulty, it can be assumed that the student is a fairly proficient speechreader. A word of cau-

tion would be to inform the youngster about the topic of conversation.

(f) The evaluator should let the youngster try out the classroom environment to see whether he or she can understand what is being said. This will provide an opportunity for him or her to hear a variety of voices and sounds in different classroom situations and determine whether he or she would be able to function in this new setting.

2. Reading

(a) Perhaps the quickest way to determine the deaf youngster's reading level is to administer a standardized reading test. However, standardized achievement tests are not designed for use with below-average students. They have only a few relatively simple items that can be answered correctly by the educationally retarded learner. Further, they do not have a cut-off point for pure chance guessing or the "orangutang score" (Frye, 1972, p. 18). A deaf youngster can earn a creditable score on a standardized reading or achievement test simply by guessing and making random responses. This is especially true when the deaf child's chronological age is equivalent to the age appropriate for the lowest grade level tested by the instrument used.

(b) In addition to standardized reading or achievement tests, the deaf youngster's reading comprehension should be tested. The evaluator can write out commands or directions that

"In conducting educational assessment of deaf children, educators typically use pictures as stimuli . . ."

require the youngster to carry out one, two, or three actions (e.g., stand up; stand up and walk to the door; stand up, walk to the door, and turn out the lights). At a more sophisticated level, the youngster can be presented with a paragraph to read followed by questions that require him or her to construct original answers, rather than the multiple-choice format typically used in standardized tests. The material used should be taken from the reading text used by the students in the class that the deaf youngster would be joining.

3. Written Language

In conducting educational assessments of deaf children, educators typically use pictures as stimuli to elicit samples of written language. These pictures may range from simple ones used at the elementary level for eliciting single sentence responses to complex ones (e.g., Norman Rockwell illustrations, captionless cartoon strips, or a series depicting a chronological series of actions). This type of picture provides the stimulus to elicit larger and more complex samples of written language. The number and complexity of the sentences obtained in the sample will, of course, depend on the youngster's vocabulary and language level and amount of time given to complete the task.

The results of an accurate and objective assessment of a deaf youngster's communication skills will help to determine whether he or she can function successfully in a regular class placement.

The assessment procedures described above might be conducted by the educational assessment services

team which is mandated by P.L. 94-142, or with assistance from a local program for the hearing impaired, a speech and hearing clinic, or a teacher of the hearing impaired.

An admissions committee should be organized by the local education agency to review the assessment findings and to make a recommendation for an educational placement that would best serve the deaf youngster's needs. The admissions committee might include the following:

1. Members of the educational assessment services team (e.g., the audiologist, psychologist, speech therapist, otologist, diagnostic teacher, and teacher of the hearing impaired).
2. The local director of special education or supervisor of services for the hearing impaired.
3. The principal of the school and the teacher of the class the child would be attending if a regular class placement were recommended.
4. The child's present teacher.
5. The child's parent(s) or guardian and the child if he or she is capable of participating in the discussion.

The evaluation report and case study to be reviewed by the admissions committee should include the following:

1. Family information
2. Birth history
3. Legal information
4. Medical history
5. Emotional adjustment evaluation
6. Auditory behavior
7. Language behavior
8. Educational history
9. Testing information and results
 - (a) Sensory
 - (b) Intelligence
 - (c) Language
 - (d) Motor
 - (e) Emotional
 - (f) Social maturity
 - (g) Educational achievement

10. Conclusions
11. Recommendations

References

- Frye, E. B. *Reading instruction for classroom and clinic*. New York: McGraw-Hill, 1972.
- Hayes, G. M., & Griffing, B. L. *A guide to the education of the deaf in the public schools of California*. Sacramento: California State Department of Education, 1967.



Psychoeducational Diagnosis of the Hearing Impaired: A Student Monitoring System

James T. Heriot
Gerard J. Maddock

Introduction

There is no single way to psychoeducationally evaluate hearing impaired persons. There are numerous ways to approximate past, current, and future functioning by integrating current methodology with a case management system unique to the needs of a particular student. Hearing impaired persons are traditionally evaluated according to a rather standard formula. The Wechsler nonverbal test batteries (Wechsler, 1974) and the Hiskey-Nebraska (Hiskey, 1966) are characteristically utilized. A recent article presents an excellent discussion of the use of the WISC with hearing impaired children (Sisco & Anderson, 1978) and relates the Wechsler battery to the Hiskey.

Both the Wechsler and Hiskey batteries stress performance-type skills and avoid analysis of linguistic functioning because it is felt that defects in verbally oriented tests would unduly penalize the hearing impaired student in terms of diagnosis and prognosis. The discussion which follows presents an approach leading to a global estimate which, in some cases, is a reflection of nonverbal abilities. However, this approach focuses on the verbal loadings of all tests in the battery without distorting estimates of the students' capabilities. This is accomplished via weighting, prorating, and multiple central tendency scores. For exam-

ple, global estimates are presented which reflect nonverbal IQ, verbal IQ, and optimized or nonstandard IQ.

The traditional global IQ approach to evaluation of hearing impairment does not, in the authors' opinion, adequately stress functional capabilities. Program casualties are the final and most costly repercussion of nonfunctionally oriented vocational/social/post-educational and linguistic evaluations. Furthermore, there is no widely accepted format to track case management/diagnostic strategies with hearing impaired persons. The authors stress *dimensional, functional* evaluation. A dimensional, functional evaluation should be united with an explicit case management strategy. This approach should be maintained from entry to exit and should lead to continuous followup.

The Student Monitoring System (SMS)

Implications and Derivations

The Arizona Diagnostic Treatment and Education Center model follows an electronic data preparation, data entry, data "message," storage, and retrieval system. Such an approach has obvious implications for clinical, research, and survey programs. The theoretical and clinical base for this model relies heavily on the works of H. J. Mark, specifically Mark (1962). The data management strategies were developed at the University of Rochester School of Medicine and Dentistry and were elaborated and implemented at the Arizona School for the Deaf and Blind (ASDB). A number of programs have adapted this approach (Heriot, 1975; Heriot & Medvene, 1975; Heriot &

Schmickel, 1967; Keller, 1978; Kotok, Kotok, & Heriot, 1977; Stebbins, Emmel, Heriot, & Rockowitz, 1975).

Suggested Model

The variety and magnitude of data generated by hearing impaired and other sensory impaired persons in the diagnostic setting necessitate an explicit, well-defined data management system to monitor and expedite student progress. This system should impose case management constraints on referral agencies, educational staff, administrative staff, followup staff, and exit agencies. In this case, "constraints" implies that all decisions made regarding a client be formal and orderly.

When a referral is made to the Arizona Diagnostic Treatment & Education Center, the client enters on Step 1 of the Student Monitoring System (SMS—Table 1). From this point on, the client proceeds along the SMS at a rate commensurate with the amount and quality of existing data. The speed with which these data can be garnered and the case-to-staff management load can lead to additional delays or accelerations in the system. If there is pre-existing data and these data meet the criteria imposed by the SMS, it is possible to "loop" through the SMS at a quicker pace. Steps in the SMS can be preempted by administrative decision or by case contingencies. Also, it should be made clear that a student can exit the system as readily as she or he enters the system.

The SMS, in addition to providing an orderly "stepped" data management and case management system responds to federal mandate (P.L. 94-142) and state mandate (Special Education Voucher Program). All teacher and administrator activities

The authors wish to express their gratitude to the following professionals at the Arizona State School for the Deaf and Blind: K. Keller, for editorial consultation, E. Babcock & R. Shingler, for editorial support, and R. Hoag, for professional and administrative support.

"... variations in learning rates ... are associated with variations in measurable cognitive functioning."

Table 1
Student Monitoring System

1. Referral & initial priority assigned	
2. Application mailed to family	
3. Application received from family	
4. Records requested from previous schools, doctors, etc.	
5. Records received	
6. Records reviewed & preliminary diagnosis established	
7. Initial home visit made	
8. Decision to accept referral made	
8a. Re-referral made	
9. Priority assigned to case	
10. Parent intake scheduled	
11. Parent intake completed	
12. Agency intake scheduled	
13. Agency intake completed	
14. Evaluation of student's present setting	
15. Inhouse/outreach status determined	
16. Voucher submitted	
17. Assignment of teacher to case	
18. Teacher reviews records	
19. Teacher views present class	
20. Diagnostic strategies established with teacher (package of instruments selected)	
21. IEP conference	
22. Entry date	
23. Psychometrics completed	} potentially concurrent
24. Audiometrics completed	
25. Ophthalmological evaluation completed	
26. Specialty evaluations completed	
27. Reassessment of preliminary diagnosis	
28. Educational plan described	
29. Educational plan implemented	
30. Educational plan evaluated	
31. Evaluation/plan discussed with family	
32. Post-exit placement determined	
33. Interagency staffing	
34. Exit plan implemented	
35. Reports sent to agency	
36. Inservice implementation of educational plan	
37. One-month follow-up	
38. Three-month follow-up	
39. Six-month follow-up	
40. Nine- to twelve-month follow-up, including second achievement testing	
41. Case placed on inactive status	

are related to and recorded by number according to the SMS. The "quick" referral, the "dial-a-test," the incomplete referral, or the inaccurate referral are precluded. Most importantly, followup and case closure are clearly specified.

Suggested Strategies

The advent of state and federal ordinances and mandates have caused the delivery of special education services to become extremely complex. An enormous amount of administrator time and paperwork are now required. A suggested order of activities for processing students for the SMS and integrating these procedures with the SMS is shown in Table 2.

Rationale

Premise

The analysis is based on the premise that abilities in some persons develop unevenly and that these variations can be isolated into specific diagnostic categories or dimensions. It also rests on the assumption that variations in learning rates as well as variations in other behavioral areas are associated with variations in measurable cognitive functioning. For persons displaying such variations, it is felt that a full-scale or global IQ (or other unitary expressions of ability) mask important strengths and weaknesses, both in terms of describing a student's current levels of functioning and in terms of future remedial planning.

Derivation of Pro-Rated IQ and the Patient Criteria Standard

In order to develop ability profiles as well as to satisfy legal needs for

"The intent . . . is to express what a child does in a familiar environment . . ."

Table 2
Referral Process

1. Receipt of referral card
2. Determine priority list number and register on referral card
3. Assign file number to case
4. Register name in alphabetical listing
5. Register name in ROSTER listing
6. Begin SMS Progress Sheet
7. Call referring agent to acknowledge referral
8. Call family to acknowledge referral and inform them of application procedure
9. Mail application to family
10. Start folder
11. If deaf-blind, place name on D-B roster and complete application
12. Place referral card in Administrative Meeting folder to present at next administrative meeting
13. Assign initial case manager at administrative meeting
14. Place card in card file
15. Place SMS Progress Sheet in notebook
16. Place file folder in student files
17. Upon receipt of completed application, mark date on Progress Sheet
18. List record requests on sheet
19. Request records from agencies (list dates requested)
20. Replace file in student files

unitary IQ statistics and various state and federal mandates, the following steps are taken with each student to establish a measure of central tendency or general ability.

1. A pro-rated IQ is established. This statistic represents the average in most cases of the established method of developing IQ estimates in sensory-impaired students, e.g., verbal measures for students with visual handicaps and nonverbal measures for students with hearing handicaps. However, the use of multiple central tendency measures allows expression of realistic limitations imposed by sensory difficulties and at the same time affords comparison to groups with the

same sensory diagnoses. A pro-rated IQ derived in this way correlates with full scale intelligence (Silverstein, 1970).

2. A Patient Criteria Standard (PCS) is derived. The PCS is a combination or average of the pro-rated IQ and an Activities of Daily Living (ADL) test. The intent of the PCS is to express what a child does in a familiar environment, using the abilities that correlate well with general intelligence. If ADL measures unduly inflate or deflate the overall measure of "Central Tendency" or the PCS, they are weighted accordingly. However, they are still taken into account clinically and educationally.

Function of the PCS

The PCS serves as a standard against which scores from all other test dimensions are compared. Scores which vary 15% or more below the PCS highlight potential deficits. The critical values or "cut-offs" are sometimes adjusted to minimize false negatives and positives in special populations, depending on the test publisher's data.

Nature of the Analysis

The analysis is dimensional, that is, it is not tied to specific test batteries. Rather, the tests selected are based on the child's mental age and are adapted to the dimensional analysis. Thus, a multivariate model is adhered to regardless of the age of the student (infancy through adulthood).

In this model, the measures of academic achievement serve as criterion variables. All other dimensions are treated as predictor variables. When the criterion variables vary significantly from peer norms, national norms, and/or the child's own norms (PCS), deficits or strengths in predictor variables are looked to as possible explanations.

Limit Testing (Optimizing)

When any dimension varies significantly (e.g., 15% or more) below the PCS, standardized limit testing or optimizing techniques are employed. These limit testing techniques are applied in the form of hypothesis-testing experiments. Typically, the experiment tests the hypothesis that the measured deficiency can be changed significantly in an upward direction.

If there are significant increases in the potential deficit due to limit testing, one can begin to formulate answers to the following questions:

"The goal . . . is to identify the major area loadings and to cluster the tests accordingly."

1. What is the most likely cause of the deficit?
2. Based on the limit testing techniques employed successfully, what prescriptive or remedial techniques should be used?
3. To what extent is the deficit real or biological vs. situationally or environmentally induced?

If in fact there is no significant change in the measured deficit following limit testing and assuming that limit testing techniques have been exhausted, one can then say with confidence that the deficit in question is real or biological rather than a situational difficulty. In order to be aware of score changes due to limit testing, they must exceed changes normally expected from series or repetition effects.

These limit testing techniques have been utilized successfully in ruling out and/or testing the impact of sensory deficits, anxiety, gross-motor coordination difficulties, linguistic problems, situational and interpersonal tester-testee variables, neuro-muscular difficulties, and a number of other parameters that can be identified as potentially affecting test performance. There is strong evidence that such procedures do not affect the validity or reliability of subsequent testing.

Variables

The variables studied cluster into four major areas. The tests do overlap across areas. The goal, however, is to identify the major area loadings and to cluster the tests accordingly. Examples of the specific tests, subtests, or items which might be used, for instance, with a school-age child, follow:

1. *Self, Parent, or Teacher Estimates of Current or Expected*

Levels of Functioning

- (a) Parental Estimate of Development Scale (PEDS, Heriot & Schmickel, 1967).
 - (b) Self-Estimate of Development (SEDS, Heriot & Maddock, 1973).
 - (c) Teacher Estimate of Development (TEDS, Heriot & Maddock, 1973).
 - (d) Vineland Social Maturity Scale.
2. *Tests Related to "g" or Full Scale Intelligence*
 - (a) Spatial relationships, e.g., WISC Block Design.
 - (b) Two-dimensional figure closure, e.g., WISC Picture Completion.
 - (c) Spoken vocabulary, e.g., WISC Vocabulary.
 - (d) Long-term retention of verbal cultural information, e.g., WISC Information.
 3. *Acquired Information*
 - (a) Academic achievement.
 1. Oral, single-word recognition, e.g., Wide Range Achievement Test (WRAT).
 2. Reading comprehension for connected material, accuracy, speed, e.g., Gates-MacGinitie test.
 3. Arithmetic calculation, e.g., WRAT.
 4. Word-to-picture reading, e.g., Peabody Pictures Vocabulary Test (Heriot, Tavormina, & Vautrain, 1973, 1974).
 5. Picture-to-word reading, e.g., Gates-MacGinitie.
 - (b) Definitional and informational tasks, e.g., WISC Vocabulary and Information (also "g" related).
 - (c) Social Maturity, e.g., Vine-

land Social Maturity or AAMD Adaptive Behavior Scales, etc.

4. Memory and Attention Substrates of Linguistics

- (a) Immediate visual and auditory recall, e.g., ITPA, BINET, WPPSI, WISC, WAIS.
- (b) Continuous visual attention or vigilance, e.g., Knox Cubes.
- (c) Learning and production of a miniature language with memory, attentional, and fine-motor components, e.g., WISC Coding.

Procedures

Tests are administered by para-professionals and/or teachers and supervised by personnel certified in the appropriate areas. Computers are used to process, edit, store, and print/retrieve results on each student. The generated report displays the results of student testing, parental intake, staff consultation, parental counseling, and, in some cases, agency staffings. Also included are supporting documents, e.g., audiometrics; visual examination results; neurologic exams; educational reports; physical therapy reports; and other specialty documents. The psychologist's comments and recommendations are unique to each child as are test results.

References

- Heriot, J. T. The psychodiagnostic technician. In R. Haggerty, K. Roghmann, & R. Pless (Eds.). *Child health and the community*. New York: John Wiley & Sons, Inc., 1975.
- Heriot, J., & Maddock, J. *Self and*

- teacher estimates of development. Unpublished manuscript, 1973.
- Heriot, J. T., & Medvene, L. Memory-attention versus perceptual motor correlates of learning rates in a pediatric population. *Pediatric Psychology*, 1975, pp. 11-17.
- Heriot, J. T., & Schmickel, C. A. Maternal estimates of intelligence in children evaluated for learning potential. *American Journal of Mental Deficiency*, 1967, 71, 920-924.
- Heriot, J. T., Tavormina, J. B., & Vautrain, R. L. The Peabody as a substitute measure of oral reading. *Perceptual and Motor Skills*, 1973, 36, (3).
- Heriot, J. T., Tavormina, J. B., & Vautrain, R. L. Correction factor for the PPVT as a reading test. *Psychological Reports/Perceptual and Motor Skills*, 1974, 38.
- Hiskey, M. S. *Hiskey-Nebraska test of learning aptitude*. Lincoln, Neb.; Union College Press, 1966.
- Keller, K. Personal communication. Arizona School for the Deaf and the Blind, Department of the Deaf, 1978.
- Kotok, D., Kotok, R., & Heriot, J. T. Cognitive evaluation of children with elevated blood lead levels. *Journal of Diseases of Children*, 1977, 131.
- Mark, H. J. Elementary thinking and the classification of behavior. *Science*, 1962, 135, 75-87.
- Silverstein, A. B. Reappraisal of the validity of the WAIS, WISC, and WPPSI short forms. *Journal of Consulting and Clinical Psychology*, 1970, 34, 12-14.
- Sisco, F. H., & Anderson, R. J. Current findings regarding the performance of deaf children on the WISC-R. *American Annals of the Deaf*, 1978, 123, 115-121.
- Stebbins, W. C., Emmel, A., Heriot, J. T., & Rockowitz, R. J. Congenital ophthalmoplegia and school achievement: A case study. *Developmental Medicine and Child Neurology*, 1975, 17, 237-243.
- Wechsler, D. *Wechsler intelligence scale for children-revised*. New York: The Psychological Corporation, 1974.

Vocational Assessment: Approaches and Implications

McCay Vernon

Any discussion of vocational diagnosis and evaluation raises the immediate question of "Diagnosis and evaluation for what?"

Although it has been well established that persons who are deaf have the same range and essentially the same distribution of intelligence and vocational assets as the nondeaf, the opportunities for deaf individuals to fulfill these potentials vary greatly depending on geographic locations. For example, in Norway, under the direction of the late Dr. Tjerle Basilear a multiply-handicapped deaf individual had excellent facilities for academic education, vocational training, and for both sheltered and competitive working and living settings. In the United States opportunities for multiply-handicapped deaf persons are not nearly so good. In fact, they are poor (Garrett & Levine, 1974, pp. 329-362).

By contrast, the United States offers college programs to bright deaf students—programs which are not available anywhere else in the world. In the same vein, it should be noted that certain countries, such as England and Germany, have demands for semi-skilled and skilled labor that are highly favorable to the employment of deaf workers, whereas in many countries this is not the case.

This brief preface is necessary in order to establish that in most geographic areas the "for what" of vocational evaluation and diagnosis is determined more by the opportunity a given area provides its deaf citizens than by the abilities of those deaf people.

This often unfortunate but practical consideration underlies the entire balance of this paper, which will outline the general principles of evaluating deaf individuals' assets and

liabilities, but which cannot address the extent to which various communities permit a deaf person the opportunities to realize these abilities.

Essentials of Effective Diagnosis

A complete diagnostic evaluation should consist of part or all of the following kinds of information:

1. A measure of intelligence.
2. An evaluation of personality structure.
3. A measure of educational achievement.
4. An appraisal of communication skills.
5. Aptitude and interest testing and evaluation.
6. Case history data.

In many cases, all these data may not be needed; or, if needed, can be obtained in part from school records or sources other than the psychological examination. For example, a diagnostic workshop involving job sampling, work experience, and more extensive observation is often a superior way to obtain full evaluations than is a formal test setting.

Before discussing each of the above six areas of evaluation separately, it is helpful to define deafness and to establish certain basic general principles for diagnosing persons who are deaf. A deaf person in the context of this paper is an individual born without sufficient hearing to understand speech or a person who loses his or her hearing before having learned language (approximately three years of age). Many of the principles to be discussed will also apply to the hard of hearing or those adventitiously deafened, but to lesser degrees. However, it is the prelingually deafened person who is likely to

be the greatest challenge to psychological assessment. Therefore, the primary focus of this paper is toward the main tools in the diagnosis of this type of individual.

General Principles for the Psychological Evaluation of Deaf Adults

Four basic concepts underlie effective psychological examinations of deaf people. These concepts form the underlying frame of reference for the discussion in this paper of the different parts of an evaluation.

1. Psychological tests that involve the use of verbal language to measure intelligence, personality, and aptitude are not generally valid because they measure the deaf person's language limitations imposed by his or her deafness. They do not measure how bright the person is or what he or she can do.
2. Tests given by persons not experienced with deaf or hard-of-hearing clients are subject to appreciably greater error than when administered by those familiar with deaf individuals.
3. Group testing with deaf and hard-of-hearing persons is a highly dubious procedure and should be regarded as only a screening technique at best.
4. Often the congenitally hard-of-hearing person is, from a psychodiagnostic point of view, much more like a congenitally deaf person than his or her speech and response to sound suggest. It is, therefore, crucial in a psychological evaluation that the individual be given tests appropriate for a person

"... a verbal test of intelligence usually does not measure the intelligence of a deaf person."

with a profound hearing loss, as well as those tests for a person with normal hearing. Where large differences appear between the sets of test responses, they often show that the individual did better on the nonlanguage tests appropriate for the deaf person. It is these findings that should be judged the more valid in such a circumstance (Sullivan & Vernon, 1979).

Intelligence Testing

One basic concept underlying the measurement of a deaf person's intelligence test, to be valid with a great majority of deaf persons, must be a nonverbal performance type instrument. Drastic and tragic consequences have resulted when verbal tests have been given to deaf people. There have been cases of above-average deaf persons being put in hospitals for the retarded (Sullivan & Vernon, 1979; Vernon & Brown, 1964), and there have been many cases of deaf people being denied proper vocational rehabilitation because inappropriate tests were employed. For example, there is at present a student in Gallaudet, the world's only college for the deaf, who spent several years in a hospital for the mentally retarded and additional years in an after-care home because his IQ, as individually measured on a verbal test, was below 50. Another girl now at the California School for the Deaf, Riverside, spent five years in a hospital for the retarded due to an "IQ" of 29 on a Binet. Retest on a non-verbal intelligence test (Leiter) yielded an IQ of 113. Another case involved a deaf man who was hospitalized as retarded at four years of age. After 35 years of hospitaliza-

tion, he escaped and went to work in Chicago. When they found him two years later, he was referred for testing to a psychologist experienced in deafness and found to have an IQ of 80. At least 15 clear-cut hospitalized cases of misdiagnosis such as these have been seen personally by the author in the last seven years.

As stated earlier, a verbal test of intelligence usually does not measure the intelligence of a deaf person. It measures the language handicap that results from his or her hearing loss. Performance IQ tests, however, do measure a deaf person's intelligence. For example, it is possible for a deaf person to score at the genius level on a performance IQ test and at the retarded level on a verbal IQ test, or even on the verbal scale of the same tests.

Verbal tests may be used in those few cases where a deaf or hard-of-hearing individual has exceptional language achievement. However, when this is not the case, it is indicative of a poor psychological evaluation to include verbal IQ scores for a deaf person, especially if these results are not carefully qualified.

Personality Evaluation

Personality evaluation of deaf individuals is a far more complex task than is IQ testing. For this reason, test findings should be carefully interpreted in light of case history data and personal experiences with the individual. In fact, it is wise for rehabilitation counselors with long experience in the field of deafness to view with skepticism results reported by examiners who are unfamiliar with deafness when these findings sharply contradict their own impres-

sions of a deaf person with whom they are working.

Due to the communication problems inherent in severe hearing loss, personality tests are more difficult to use with deaf subjects than with the general population (Graham & Kendall, 1960; Levine, 1960, pp. 225-226; Myklebust, 1954, p. 313; Sullivan & Vernon, 1979). Not only do these tests depend on extensive verbal interchange or reading skill, but they also presuppose a rapport and confidence on the part of the subject that is difficult to achieve when the person examined may not fully understand what is being said or written. Paper and pencil personality measures are perhaps suitable for hearing impaired persons with well-developed expressive and receptive language skills, but such individuals are the exception; even the problems of test administration and interpretation make the meaningfulness of results highly suspect (Graham & Kendall, 1960; Levine, 1960, pp. 225-226; Myklebust, 1954, pp. 121-122). If projective measures like the Rorschach or Thematic Apperception Tests are used, fluency in manual communication (the language of signs) by the examiner is mandatory.

There is some question as to whether the norms for the personality structure of hearing people are appropriate for deaf and hard-of-hearing subjects (Sullivan & Vernon, 1979; Zeckel, 1934, 1942). Conceivably, deafness alters the perceived environment sufficiently to bring about an essentially different organization of personality in which normality would then differ from what it is in the case of a person with normal hearing. Although this is presently an unresolved problem, it is one that is frequently raised by scholars in the field of deafness and

"Many extremely bright and capable deaf people lack the ability to speak and lipread."

should be considered in any discussion of the personality of those with severe hearing loss.

The use of interpreters who express the psychologist's directions in fingerspelling and the language of signs is a questionable procedure. What is required is an interpreter, not only fluent in manual communication, but also knowledgeable in psychology and testing (Sullivan & Vernon, 1979). Obviously, such individuals would be doing the examining themselves and not interpreting it for others. Therefore, results reported where an interpreter is involved are not likely to meet standards of validity (Vernon, 1967).

In connection with personality evaluation, it is important to note that the confusion and disassociation reflected in the writing of deaf persons with low-level verbal skills rarely indicate an equally deranged thought process. It is usually only the result of language deficiency. Psychologists unaware of this have been known to equate the written language of semiliterate deaf persons with that of the schizophrenic, and unfortunate diagnoses have been made based in large part upon this confusion.

Few personality tests have had wide successful application with deaf or hard-of-hearing adults because of the difficulties that have been pointed out above. Three of the more commonly and effectively used are the Draw-A-Person, House Tree Person, and Bender-Gestalt.

Educational Achievement

In many cases, the referral source will want a measure of the level of educational achievement of a deaf individual.

The important and difficult aspects of achievement testing are (a) choosing a test at the appropriate educational level, (b) making sure directions are understood, and (c) interpreting results in terms of norms appropriate to deaf youth. For example, research in the United States indicates that only about 5% of graduates from day and residential schools for the deaf attain a 10th-grade level in educational achievement; 41% attain seventh or eighth grade; 27% reach fifth- or sixth-grade level; and approximately 30% are in fourth grade or below. Most of the last category are termed functionally illiterate (McClure, 1966).

These figures are given in order that psychologists will understand some of the educational problems of deafness and have some baseline data about the norms to be used in judging the individual educational achievement findings of a particular client.

An Evaluation of Communication Skills

As it is in the realm of communication that deafness presents its major handicap, it is important that an evaluation include an assessment of communication skills.

There are three aspects of communication that should be appraised in a deaf individual. The first and by far the most important is the ability to read and write. It has been found in research that this is the modality most widely used in job settings by persons who are deaf (Vernon, 1970). Therefore, it is obviously important to determine how effective the individual is in exchanging ideas in written language. The person's degree of skill will go a long way toward deter-

mining the type and level of occupational skills open to him or her.

Speech and speechreading are the other key aspects of communication to be evaluated. These skills have considerable potential value to a deaf person in the world of work. To be able to speak intelligibly is especially important and helpful. Speechreading ability is an asset, but even the most skilled deaf lipreader generally finds this an inadequate way to obtain important information and will usually prefer writing. When one remembers that 40-60% of the sounds of English look just like some other sound on the lips, it is understandable that the confusion of speechreading would lead a deaf person to the frequent use of writing.

One final point must be made clear in evaluating communication skills. Many extremely bright and capable deaf people lack the ability to speak and lipread. It is of critical importance that rehabilitation specialists not confuse difficulty in communication with lack of intelligence. It is also possible that a congenitally deaf person may have a very high IQ yet have written language that is not grammatically sophisticated. Rehabilitation specialists must be keenly aware of these factors if they are to be fair and helpful to a deaf client.

Aptitude and Interest Testing

A basic part of a complete psychological evaluation is aptitude testing, i.e., finding out the particular abilities that an individual may have. As there are hundreds of tests for this on the market, it is not feasible to list or discuss them individually. However, certain information about the following three general areas of aptitude may be of great value with a

"There are pictorial tests designed for use with people who are deaf, but they are narrow in scope . . ."

deaf person because these kinds of aptitudes are directly related to the types of work done by most deaf people (Schein & Delk, 1974, pp. 12-22).

1. Manual dexterity.
2. Mechanical aptitudes.
3. Spatial relations.

It is important in selecting from the many measures of aptitude available to choose those tests that are not primarily dependent on language for either their directions or administration.

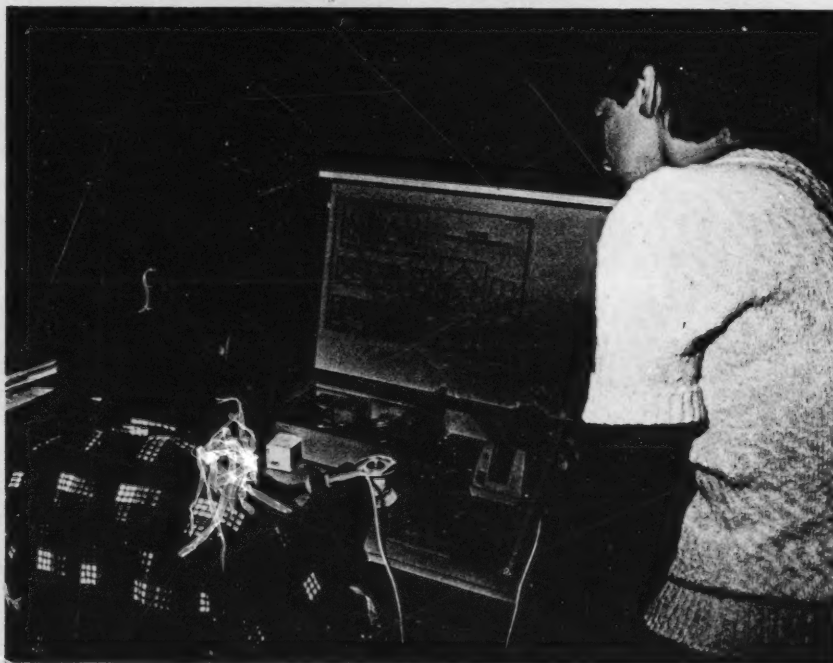
Interest tests are almost without exception highly verbal and, therefore, cannot generally be used effectively with a deaf person. There are pictorial tests designed for use with people who are deaf, but they are narrow in scope and offer limited data for a psychological evaluation.

Case History Data

The past is still the one best predictor of the future. For this reason, complete background information is of extreme importance, especially if the person is deaf and may not be accurately evaluated with regular psychological procedures. An indication of the importance of case history data is the fact that the best psychiatric and psychological evaluations often are based 75% upon background information.

Some factors to consider are the person's performance in past jobs, whether or not he or she has habitually demonstrated particular problems or assets, what kinds of circumstances have led to the person's success or failure and what specific educational and vocational skills have been mastered.

In the case of young deaf clients from residential or large day schools,



"... the opportunities open to a deaf individual should be the overriding factor in determining what diagnostic devices will be used ..."

it is often possible to obtain complete and insightful information by contacting their school.

Summary

In sum, the opportunities open to a deaf individual should be the overriding factor in determining what diagnostic devices will be used with that person. It does little good to diagnose architectural ability if the deaf individual is not going to be given a chance to be an architect.

Within this framework, the main diagnostic tools are the psychodiagnostic instruments for intelligence and personality assessment, educational achievement tests, appraisal of communication skills, evaluation of aptitude and interests, and case history data. In certain cases, some of these measures are best obtained in an evaluation workshop and/or through job sampling.

References

- Garrett, J. F., & Levine, E. S. *Rehabilitation practices with the physically disabled*. New York: Columbia Press, 1974.
- Graham, F. D., & Kendall, B. S. *Memory-For-Designs Test: Revised general manual*. *Perceptual Motor Skills*, 1960, 2; 7.
- Levine, E. S. *The psychology of deafness*. New York: Columbia University Press, 1960.
- McClure, W. J. Current problems and trends in the education of the deaf. *Deaf American*, 1966, pp. 8-14.
- Myklebust, H. *Auditory disorders in children*. New York: Grune & Stratton, 1954.
- Schein, J. D., & Delk, M. T. *The deaf population of the United States*. Silver Spring, Md.: National Association of the Deaf, 1975.
- Sullivan, P. M., & Vernon, M. Psychological assessment of hearing impaired children. *The School Psychology Digest*, 1979, 8, 271-290.
- Vernon, M. A guide to the psychological evaluation of deaf and severely hard of hearing adults. *Deaf American*, 1967, 19, 15-18.
- Vernon, M. Potential achievement and the rehabilitation of the deaf. *Rehabilitation Literature*, 1970, 31, 258-267.
- Vernon, M., & Brown, D. W. A guide to psychological tests and testing procedures in the evaluation of deaf and hard of hearing children. *Journal of Speech and Hearing Disorders*, 1964, 29, 414-423.
- Zeckel, A. Research possibilities with the deaf. *American Annals of the Deaf*, 1942, 87, 173-191.
- Zeckel, A., & Kalb, J. J. A comparative test of groups of children born deaf and of good hearing by means of the porteous test. *American Annals of the Deaf*, 1934, 84, 114-123.

Assessing Communication Skills

Anne B. Spragins
Dennis R. Cokely

A comprehensive assessment of a deaf child is a multifactored procedure involving professionals in education, psychology, audiology, speech, and other related areas. At the heart of the procedure is the child's way of communicating with the world. Deaf children communicate in various ways that are not immediately obvious to the evaluator new to deafness. One short article cannot make the uninitiated aware of all the subtleties of communication among the deaf community. It can, however, raise issues for which there are many questions, if not many answers.

A valid comprehensive assessment of a deaf child's skills and acquisitions depends on conducting the evaluation procedures in the mode of communication normally used by the child. According to the now familiar rules and regulations of Public Law 94-142, all assessment procedures must take place in the child's native language. The difficulty for the evaluator is in determining the "native language" for an individual deaf child. The regulations for PL 94-142 read:

Section 121a.9 Native Language *Comment.*

- (1) In all direct contact with a child (including evaluation of the child), communication would be in language normally used by the child and not that of the parents, if there is a difference between the two.
- (2) If a person is deaf or blind, or has no written language the mode of communication would be that normally used by the person (such as sign language, braille, or oral communication).

(Federal Register, 1977, p. 42479)

Clearly, the intent of the law is to suggest that failure to determine the child's "normally used" mode of communication could result in invalid or legally vulnerable evaluation results. Before proceeding with further assessments, evaluators are forced to assess communication skills themselves or to gather reliable information on the deaf child's communication mode. How might an evaluator proceed? A first step is to become aware of several factors affecting the type of communication methods used by deaf children.

Factors Affecting Communication Skills

Two major factors that will affect the child's communication skills are (a) the impact of communication methodologies in deaf education and (b) the sociolinguistic situation of the deaf child and the deaf community. Communication methodologies used in deaf education attempt to develop the English language competencies of the deaf child by either of two approaches: (a) emphasis on development of oral/aural skills or (b) utilization of simultaneous communication (i.e., producing a manual representation for English while speaking or mouthing English). The methodology used with the child in the classroom may have an effect upon, but will not necessarily determine the communication mode normally used by the child.

The evaluator cannot assume that the methodology used in the child's current educational placement is the child's "normally used" mode of communication. The child may have had a varied communication/education background. The question for the evaluator is: "What has been the

impact of the educational experience for that particular child?"

A second consideration in determining the child's preferred mode of communication is the child's sociolinguistic situation. A number of deaf children reach school age with linguistic competence in a language other than English—American Sign Language (ASL). These children, usually deaf children of deaf parents, comprise about 10% of the hearing impaired student population. Evaluators who have limited signing skills or a monolingual attitude toward linguistic differences may fail to capitalize upon these children's existing linguistic skills. As a result, many of these children have, in the past, been seen as linguistically retarded and have been placed in inappropriate educational settings. Evaluators are urged to review carefully the total language environment of the deaf child before reaching hasty conclusions about linguistic competence.

Determining the deaf child's normally used mode of communication can be a difficult task. At present, there are no quantifiable screening devices to determine whether the child's most comfortable means of communication is ASL, a manual representation of English, or speech. The evaluator is forced to rely upon information from the parent, the teacher, the child, and observations. In addition, whenever possible, the evaluator should employ the services of a deaf adult who is proficient in ASL to provide observational data.

A thorough evaluation of the communication skills of a deaf child would include assessment in those areas specified in Table 1.

Few evaluators or teams of evaluators could provide a thorough communication assessment in each of the areas listed. Neither assessment

“... the evaluator must rely heavily upon the subjective judgment of teachers, parents, and the students themselves.”

Table 1

Communication Assessment			
ASL		English	
receptive	expressive	receptive	expressive
		manual codes	manual codes
		audition	speech
		speechreading	writing
		reading	

Table 2

Behavioral Description of Speech Intelligibility Ratings

Rating	Behavioral description
Very intelligible	Similar to the speech of a hearing person of the same age
Intelligible	Somewhat difficult to understand.
Barely intelligible	Can only understand after repetition and use of other clues.
Not intelligible	—
Student would not ordinarily attempt to use speech	—

techniques nor trained evaluators are available. However, deaf children must be evaluated for appropriate educational placement. The following discussion may serve as a guideline for determining the child's preferred mode of communication.

Components of Communication Assessment

Sign Skills

An initial question is whether the child possesses competence in ASL or some manual representation for English. Evaluators should be aware

that a child who signs is not necessarily using ASL. Significant differences between ASL and manual representations for English exist on the lexical, syntactic, and semantic levels. Descriptions of some of the major characteristics of ASL (Baker & Cokely, in press) as well as descriptions of some of the characteristics of manual representations of English (Bornstein, 1973; Wilbur, 1976) should prove helpful. Evaluators need to be familiar with these features as they may need to be quite specific about which particular code for representing English the child is comfortable using.

It is tempting to consider using tests designed to assess speaking/listening skills in English to measure a student's language competence. The evaluator simply signs the stimuli and considers this a valid measure of a student's competence. However, because these tests were developed for assessing English skills (and not ASL or a coded form of English), a built-in bias exists which may affect the test results when the tests are translated into signs. Currently, no standardized test exists which is appropriate for assessing student competence in ASL or in a manual representation of English. Given this situation, the evaluator must rely heavily upon the subjective judgment of teachers, parents, and the students themselves. It is here that the evaluator can use the services of a deaf adult who is fluent in ASL and familiar with the various manual codes for English. The inclusion of such an observer should help insure protection of the deaf child's rights to a free and appropriate education.

Speech Intelligibility

The assessment of speech intelligibility may be a factor in determining the appropriateness of placement in a mainstream setting. The assessment is a subjective judgment of the evaluator, commonly expressed in terms of relative intelligibility. Assessment by or consultation with the speech pathologist also can be of great value. A behavioral description developed by Jensema, Karchmer, and Trybus (1978) for research purposes could be used to help the evaluator develop an internal norm. (See Table 2.)

Results of the study by Jensema et al. (1978) on hearing impaired children aged 4-23 years indicated that speech intelligibility did not im-

"... the underlying variables for whether the child speaks will be ... the level of hearing loss."

prove with age, but was more strongly related to degree of hearing loss. In addition, when degree of hearing loss was taken into account, speech intelligibility did not vary in relation to type of educational placement (residential school, day school, full-time class, part-time service). These results are consistent with the findings from a longitudinal study (Babbini & Quigley, 1970) of 163 students in six residential schools for the deaf. Data from this study indicated that the underlying variable for whether the hearing impaired child speaks will be (given today's practices in teaching speech) the level of hearing loss. It is important that the level of speech intelligibility not be confused with unrelated variables (e.g., level of mental functioning). The inability to speak, for example, does not imply a lack of intelligence or cognitive development.

Speechreading

Speechreading or lipreading ability of hearing impaired children may be a source of misunderstanding for people new to deafness. Only 33% of the isolated English sounds are readily visible on the lips (Hardy, 1970). Many elements that sound different appear the same on the lips (e.g., bat, mat, pat). In speechreading, the reader must use situational variables and knowledge of the grammar of a language to decode messages. The prelingually deaf child (deafened before language development) who has no foundation of oral language on which to depend may find it difficult to become a good speech reader.

Evaluation of speechreading ability has been conducted with the use of several instruments. These include the *Craig Lipreading Inventory* (Craig, 1964), the *Diagnostic Test of Speechreading* (Myklebust & Ney-

bus, 1971), and the *Utley Test of Lipreading* (Utley, 1946).

Summary

Comprehensive evaluation of the deaf child may depend on determining the child's preferred mode of communication; however, making that determination is a complex problem that requires skill and training. In the case of the deaf child, professionals need to proceed with caution, with care, and with a willingness to learn from those who are deaf themselves.

References

- Babbini, B., & Quigley, S. *A study of growth patterns in language communication and educational achievement in six residential schools for deaf students*. Urbana, Ill.: Institute for Research on Exceptional Children, Illinois University, 1970.
- Baker, C., & Cokely, D. *American Sign Language: A teacher's resource text on grammar and culture*. Silver Spring, Md.: T. J. Publishers, Inc., in press.
- Bornstein, H. A description of some current sign systems designed to represent English. *American Annals of the Deaf*, 1973, 118, 454-470.
- Craig, W. Effects of preschool training on the development of reading and lip-reading skills of deaf children. *American Annals of the Deaf*, 1964, 109, 280-296.
- Hardy, M. Speechreading. In H. Davis & S. Silverman (Eds.), *Hearing and deafness*. New York: Holt, Rinehart and Winston, 1970.
- Jensema, C., Karchmer, M., & Trybus, R. *The rated speech intelligibility of hearing impaired children: Basic relationship and a detailed analysis*. Series R., Vol. 6. Washington, D.C.: Office of Demographic Studies, Gallaudet College, 1978.
- Myklebust, H., & Neyhus, A. I. *Diagnostic art of speechreading*. New York: Grune & Stratton, 1971.
- Utley, J. A test of lipreading ability. *Journal of Speech Disorders*, 1946, 11, 109-116.
- Wilbur, R. The linguistics of manual languages and manual systems. In L. Lloyd (Ed.), *Communication assessment and intervention strategies*. Baltimore: University Park Press, 1976.

Audiological Evaluation

Antonia B. Maxon

In order to determine appropriate services for individuals with hearing impairment, professionals in education and related fields need some understanding of various audiological tests and their interpretation. The optimal auditory management of a hearing impaired individual requires, as a first step, a comprehensive diagnostic description of his or her particular hearing loss.

Types of Hearing Impairment

There are three commonly distinguished types of hearing impairment: conductive, sensori-neural, and mixed. They can be distinguished in terms of the type of results obtained through an audiological evaluation and according to the area of the ear that is damaged. A *conductive* hearing impairment is due to damage (structural or by infection) in the external and/or middle ear. Such an impairment is considered reversible through medication and/or surgery. The accompanying hearing loss occurs because of a problem in sound transmission through the external or middle ear. A *sensori-neural* impairment is due to damage in the inner ear (cochlea) and/or peripheral portion of the VIIIth Cranial Nerve (auditory). Such an impairment is generally not considered to be medically reversible but is usually remediated through the use of amplification (hearing aids, wireless FM systems, auditory trainers) and aural rehabilitation. The accompanying hearing loss occurs because of a problem in the transmission of sound within the cochlea or along the auditory nerve. A *mixed* hearing impairment is due to damage in the external and/or middle ear and concurrent damage

to the cochlea, i.e., the impairment is both conductive and sensori-neural in nature. Determining the type of hearing loss is accomplished through pure tone audiometric testing (air conduction vs. bone conduction thresholds), electroacoustic impedance measurements, and otological examination.

Audiological Equipment

The basic piece of equipment used by the audiologist in evaluating hearing is the audiometer. It allows for precise control of the various stimuli available. One component, the pure tone generator, produces frequencies calibrated in Hertz (Hz) or cycles per second (cps). These frequencies range from 125 Hz to 8000 Hz. The intensity of the pure tone produced is controlled through attenuators and amplifiers and is governed by the intensity dial, calibrated in decibels (dB).

The intensity dial on the audiometer is calibrated in dB HTL (hearing threshold level) with normal human hearing as the referent. When normal hearing of humans is measured in dB SPL (sound pressure level) i.e., using physical pressure of sound as a basis (see Table 1), thresholds are found to be frequency dependent, that is, hearing is most sensitive in the mid-frequency range, least sensitive in the low-frequency range, and sensitivity in the high frequency range falls between those two. Since it is easier to see a deviation from a straight line rather than a curve, the normal hearing SPL curve was changed into a straight line by using a different decibel notation, i.e., using dB hearing threshold level (HTL). In this notation, OdB HTL equals the threshold in dB

SPL found at each individual frequency. This relationship of dB SPL to dB HTL can be seen in Table 1.

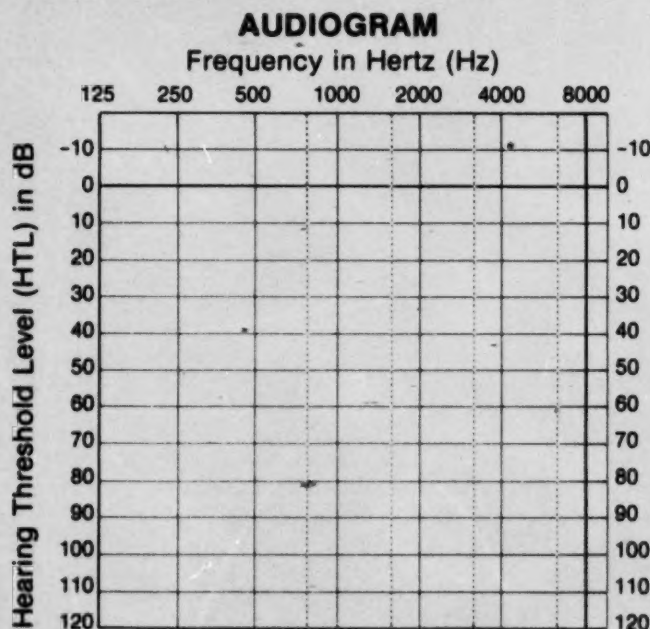
A clinical audiometer calibrated in dB HTL has the normal hearing SPL curve built into it so that when the intensity dial is set to OdB HTL the audiometer automatically produces the correct number of dB SPL for that frequency. This calibration is also valid for the higher intensities.

The audiometer is designed also to produce the signals available through several different presentation modes. One mode is air conduction (AC), routing the signal through headphones to either the right or left ear, allowing for an assessment of the entire peripheral mechanism (the external, middle, and inner ear). A bone conduction (BC) signal is produced through the bone transducer, a mechanism that vibrates, transmits these vibrations to the bones of the skull, and so to the cochlea. During testing, the transducer is placed either behind the ear (mastoid proc-

Table 1
Normal hearing thresholds in dB
SPL vs. dB HL

Frequency (Hz)	Average threshold dB SPL (ANSI, 1969)	Average threshold dB HL
125	45.0	0
250	25.5	0
500	11.5	0
1000	7.0	0
1500	6.5	0
2000	9.0	0
3000	10.0	0
4000	9.5	0
6000	15.5	0
8000	13.0	0

"Sound field assessment can also be made through the use of loudspeakers."



	RIGHT	LEFT
Air Unmasked	○	×
Air Masked	△	□
Bone Unmasked	<	>
Bone Masked	┌	┐
No Response	↙	↘
Sound Field	\$	
CNT	Could Not Test	
DNT	Did Not Test	

Figure 1
Pure tone audiogram with symbol key.

ess) or on the forehead, and only the cochlea's condition is assessed because the sound has bypassed the external and middle ear. It is important in a routine audiological evaluation to test both by AC and BC, in order to determine the type as well as the degree of hearing loss. Sound field assessment can also be made through the use of loudspeakers. Such an arrangement becomes necessary when headphones cannot be used, e.g., with very young children or when responses from hearing aid wearers are desired.

Speech stimuli can be presented either as "live voice" or in prerecorded format. In the "live voice" presentation, a microphone is used, and the audiologist monitors the intensity of his or her voice through a meter on the control panel. With prerecorded stimuli, a signal from either a turntable or a tape recorder is fed through the audiometer. In either case, the intensity of the signal is controlled by manipulating the intensity dial. A talkback system, a microphone in the room with the client and a small loudspeaker in the room with the audiologist, allows for communication between the two sections of the soundproof booth.

Aside from producing pure tones and speech, the audiometer is capable of producing various types of noise: white noise (a broad-frequency range), narrow band noise (a narrow-frequency range), and speech noise. The use of these noise stimuli (masking) is employed when it is necessary to assure that the nontest ear is not responding. The audiometer is designed so that the test stimuli can be delivered through one headphone or the bone transducer and the masking noise through the opposite headphone.

"The relationship between frequency and intensity indicates the degree of hearing loss . . ."

All of the equipment described should be in a double-chamber soundproof booth with the audiometer on one side, the headphones and loudspeakers on the other, and a soundproof wall and window between.

Figure 1 illustrates a typical audiogram form used in a clinical setting. Intensity in dB HTL is displayed along the vertical axis, with the lowest intensity (indicating the least degree of hearing loss) at the top and the highest intensity (indicating the greatest degree of hearing loss) at the bottom. Frequency in Hz is displayed along the horizontal axis with low frequencies at the left side and high frequencies at the right. The relationship between frequency and intensity indicates the degree of hearing loss in a particular frequency region. The symbols displayed in the key in Figure 1 are typical of those used in most clinics, e.g., O = right AC thresholds, X = left AC thresholds. From the audiogram, the degree (relative to intensity) as well as configuration (relative to frequency) of the hearing impairment can be determined. The type of hearing impairment can also be determined by comparing the intensity levels at which the AC and BC thresholds were obtained.

Conductive: BC thresholds are within normal limits, and AC thresholds are poorer than normal, i.e., an "air-bone gap" exists (see Figure 2.).

Sensori-neural: BC and AC thresholds are equal to each other, and both are poorer than normal (see Figure 3).

Mixed: BC thresholds are better than AC thresholds but both are poorer than normal, i.e., an "air-bone gap" exists (see Figure 4).

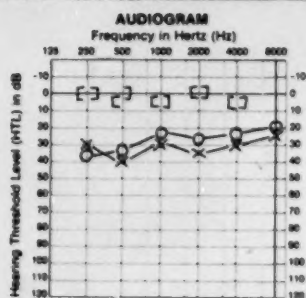


Figure 2
Mild conductive hearing loss in both ears.

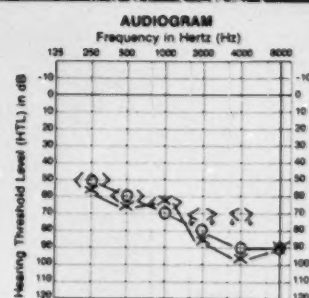


Figure 3
Moderate to profound sensori-neural hearing loss in both ears.

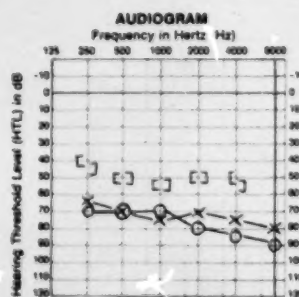


Figure 4
Severe to profound mixed hearing loss in both ears.

The degree of hearing loss can be categorized in the following manner:

Category	Threshold
Normal	10-25dB HTL
Mild	26-40dB HTL
Moderate	41-55dB HTL
Moderately severe	56-70dB HTL
Severe	71-90dB HTL
Profound	90+dB HTL

It should be noted that these categories are based only on pure tone averages (500, 1000, 2000Hz average) and do not indicate the degree of hearing loss in the higher frequencies, which must also be considered.

Pure Tone Audiometry

The purpose of pure tone audiometry is to determine a person's thresholds, the lowest intensity level at which 50% of the presented signals are heard. Since clear instructions need to be given, profoundly hearing impaired individuals or other people with language difficulties may need a demonstration of the procedures. In general, the adult and older child will be able to raise his or her hand in response to the signal but for younger children, variations in response mode will be necessary.

Generally, individual ear AC thresholds are obtained first starting at 1000 Hz. Once a response is obtained, the signal is increased by 20dB to be sure the person knows how the signal sounds, and the threshold finding procedure begins. After testing at 1000 Hz, thresholds are found for the high frequencies (2000, 4000, 8000 Hz) and then the low frequencies (250, 500 Hz). Once all of the AC thresholds have been obtained, it is necessary to find BC thresholds with the bone transducer

placed on the mastoid process of the better ear. All of the same frequencies are tested except 8000 Hz because the bone transducer cannot produce such a high frequency signal.

Once the AC and BC thresholds are obtained, it must be determined whether or not masking is necessary. When a signal is presented to the test ear over the headphones, a certain amount of that sound may be transmitted to the nontest ear. If the test signal is loud enough it will vibrate the bones of the skull and be directed to the opposite cochlea (nontest ear). The skull itself blocks a certain amount of sound from crossing over from the test ear to the nontest ear. The blockage of cross-over is referred to as interaural attenuation and is clinically considered to be approximately 40dB for AC. For BC there is 0dB interaural attenuation, i.e., any sound delivered to one cochlea can be heard by the opposite cochlea. When testing by bone conduction, it is very difficult to determine which ear has actually heard the signal unless masking is used. In order to assure that the nontest ear does not hear the test signal, one has to present noise to or "mask" the nontest ear.

Speech Audiometry

Speech audiometry is used to determine how well a person responds to speech at threshold and supra-threshold levels. Typically, the first speech measure taken is the *Speech Reception Threshold* (SRT), measured in dB, which should agree within 5dB with the pure tone speech frequency (500, 1000, 2000 Hz) thresholds. The stimuli used to obtain an SRT are spondaic words, two-syllable words with equal stress on each

syllable. These words are heavily vowel-loaded and can often be understood from correctly perceiving the vowel information only. The procedure for obtaining an SRT is to start presenting the words at 40dB above the *pure tone average* (PTA). The person repeats the words while the intensity level is decreased in a step-wise fashion. The SRT is the lowest intensity level at which three out of six words are correctly repeated. If each SRT does not match the PTA, it may be diagnostically significant and may indicate an unreliable test, an emotional problem, a language disorder, or malingering.

Discrimination of speech is assessed at a level at which the person can comfortably hear the stimuli. The words used are monosyllabic in lists of 50 and should be within the vocabulary of the person and contain a variety of speech sounds. In the testing procedure, the full list of words, e.g., W-22 lists, is presented with a carrier phrase, e.g., "Say the word _____" at 30-40dB above the SRT for each ear. The person must repeat the entire word correctly in order to be given credit for a correct response. The percent correct discrimination score (2% for each word) should agree with the configuration of the audiogram; the poorer the hearing in the high frequencies, the poorer the score. This relationship between high-frequency hearing loss and speech discrimination occurs because good discrimination is dependent on perceiving the high-frequency consonant information correctly. If the discrimination score does not match the audiogram configuration, it may be diagnostically significant, e.g., discrimination scores significantly poorer than expected may indicate a tumor on the auditory nerve. Other than the W-22 and

"The responses the audiologist expects to obtain depend on the age of the infant."

other monosyllabic word lists, there are closed-set and sentence materials available for clinical assessment of speech discrimination.

Although the SRT and speech discrimination scores are the speech measures usually obtained in a routine evaluation, others are also important. With the person who cannot understand and repeat the spondee words even at the output limits of the audiometer, it may become necessary to obtain a *Speech Awareness Threshold* (SAT). Such a measure gives an indication of the intensity level at which the person can just detect that he or she heard speech. Since the SAT is approximately 10dB better than the SRT, the latter can be predicted from the SAT.

The level at which the person finds speech to be most comfortable, *Most Comfortable Loudness* (MCL), is important in the prescription of a hearing aid because it gives an indication of a comfortable amount of amplification. Similarly, the *Tolerance Level*, the level at which speech is too loud, is also important in prescribing a hearing aid. A hearing aid that amplifies to a level higher than the tolerance level will be uncomfortable and may lead to reduced utilization of the aid.

Impedance Measurements

Impedance measurements help to determine the type of hearing loss, i.e., is it pure sensori-neural or is there some conductive component to the loss. This is done by assessing the middle ear function through the evaluation of the eardrum (tympanic membrane) mobility, middle ear pressure, presence of fluid in the middle ear cavity, and middle ear (stapedial) muscle functioning.

The Pediatric Patient

Many of the techniques described earlier must be modified in order to evaluate the hearing of infants and young children. Basically, the techniques can be separated into those used with infants (0-2 years) and those used with the young child (2-5 years).

The infant is usually tested in the sound field setting, i.e., in a sound proof booth with the stimulus presented through loudspeakers. When testing in the sound field, the responses obtained are those of the better ear so that if a difference between the ears exists, the minimal response levels obtained are those of the better ear. It is important to maintain quiet and restrict distracting movements so that the child's attention can be attracted by the auditory stimuli. Generally, the parent holds the child facing the audiologist through the window separating the test rooms. A trained observer in the room with the parent and child reports whether or not the child has responded to the sound presented. Generally, the first stimuli which are presented are loud (70-90dB HL). When a response is obtained, the intensity is decreased until the minimum intensity level at which the child will give a consistent reliable response is found. It should be noted that these are not considered thresholds, but rather minimum awareness levels with thresholds being somewhat better.

The stimuli used with infants have to be interesting so it is often suggested that noisemakers such as rattles, bells, and squeak toys be selected to represent various approximate frequency ranges. Speech may also be a good stimulus where the audiologist imitates the child's vocal-

ization, calls the child's name, or makes crying or kissing sounds. Often music, particularly that which is "up tempo" is something to which children respond easily. For the older infant, good responses may be obtained to warbled pure tones and pulsed narrow band noise, therefore, a more accurate estimate of the child's hearing at various frequencies can be made.

The responses the audiologist expects to obtain depend on the age of the infant. In general, responses such as eye widening, eye blinking, arousal from sleep, startle, cessation of crying, cessation of activity, and starting to cry are both good and reliable. Prior to seven months of age, very rudimentary localization may be obtained as a response but after that age direct localization to a sound is a good consistent response.

Aside from the distraction type of techniques described above, *visual reinforcement audiometry* (VRA) can be used with the older (12-30 months) infant. The *conditioned orienting reflex* (COR) audiometric technique is the most commonly used of the VRA procedures. An auditory stimulus is presented simultaneously with a visual stimulus (usually a toy which lights up), and the child is conditioned to turn to look to the visual display whenever he or she hears the auditory stimulus. Once the child has been conditioned, the visual display is used to reinforce the correct response to the auditory stimulus. Consistent reliable localization responses can readily be obtained with this procedure.

Impedance techniques also can be used with infants. The most difficult aspect is keeping the child relatively still so that exact measurements can be made.

Testing young children (2-5

"The audiologist says a word and the child must point to the correct picture . . ."

years) requires that different modifications be made. For these children, the headphones can usually be used and individual ear information obtained. To obtain pure tone thresholds, a play audiometry technique is used which consists of having the child make a motor "play" response upon hearing a pure tone. Whatever activity is chosen, it is helpful to have a trained assistant in the booth with the child to reinforce a correct response, say "no" to an incorrect response, and to keep the child's attention to the task.

Once the child has been conditioned, a complete AC and BC audiogram can be obtained. If the child seems to be tiring of the testing, the audiologist may choose to obtain speech audiometry information before BC, assuming that information about the middle ear can be obtained through impedance measurements. If it is necessary to use masking to assure that the responses are those of the test ear, it can easily be carried out with play audiometry.

With most of the children on whom it is possible to get a complete audiogram, it is also possible to carry out speech audiometry. The SRT can be obtained by using "spondee" toys or pictures, i.e., items which represent spondee words with which the child is familiar. The child's task is to point to the picture or toy which the audiologist has named. Speech discrimination scores can be obtained through procedures, depending on the age of the child. For the older child, an adult procedure can be used with special lists of words within the vocabulary of the average kindergarten child (PB-K lists). This procedure may be difficult or impossible to use with the child who has limited language or very poor articulation. In such instances, a picture

pointing task, such as the *Word Intelligibility by Picture Identification* (WIPI) may be helpful. The audiologist says a word and the child must point to the correct picture out of six on each of 25 plates. For the child with a severe to profound sensorineural hearing loss, a purely auditory discrimination score may not be useful if the child relies heavily on visual cues. In such cases, two different discrimination scores can be obtained, one purely auditory and one in a combined auditory and visual mode.

For the profoundly hearing impaired child who cannot accomplish either of these tasks, Ling's "five speech sounds" can be used to quickly check whether he or she is able to hear certain speech sounds. The five sounds, /a/ as in "ma," /u/ as in "shoe," /i/ as in "me," /s/ as in "sit," and /sh/ as in "she" are presented at a comfortable listening level, and the child just has to indicate that a sound is heard. This technique will give the audiologist information as to which speech sounds the child hears, but will not tell whether or not he or she can discriminate among these or any other sounds.

Screening

The purpose of a screening program is to identify children with a hearing loss so that amplification or other recommended treatment can be started and suggestions can be made for habilitation and educational placement. The earlier the identification, the better the chances of good speech and language development. The Joint Committee on Infant Screening recommends screening those infants considered a high risk for hearing impairment. This

screening usually takes place in the hospital. Those infants who do not pass are referred to an audiologist for further testing.

The American Speech-Language-Hearing Association (ASHA) has developed screening guidelines to identify those preschool and school-age children with an educationally handicapping hearing impairment. Pure tone screening is carried out according to the following procedures:

1. Test frequencies shall be 1000, 2000, and 4000 Hz.
2. Screening levels shall be 20dB HTL at 1000 Hz and 25dB HTL at 4000 Hz.
3. Failure criterion shall be failure to respond at the recommended screening level at one or more frequencies in either ear.
4. Mandatory rescreening for all failure shall occur within one week.
5. Audiologic evaluations are to be accomplished for individuals failing the screening.
6. Medical referrals should be based on the results of audiologic evaluations.
7. Screening personnel should have professional preparation in screening young children. (ASHA, 1976)

Although not recommended in these guidelines, it is good practice to include screening at 500 and 6000 Hz in order to assure that those children with conductive or high-frequency hearing impairments are identified.

Once a child or an adult has been identified as having a hearing impairment and the type and degree of the hearing impairment have been determined, the rehabilitation procedure must be started. Recommendations

must be made for personal and/or school-worn amplification, management, and educational placement. The more people working with the hearing impaired child know about the ways in which assessment is done and its implications, the better able they will be to help in the education of the child.

Bibliography

- Berlin, C. I. *Programmed instruction on the decibel in clinical audiology*. Louisiana State University School of Medicine, 1970.
- Bess, F. H. *Childhood deafness: Causation, assessment and management*. New York: Grune and Stratton, 1977.
- Davis, H., & Silverman, S. R., *Hearing and deafness*. New York: Holt, Rhinehart and Winston, 1970.
- Jerger, J. *Handbook of clinical impedance audiometry*. New York: American Electromedics Corporation, 1975.
- Katz, J. *Handbook of clinical audiology*. Baltimore: Williams and Wilkins, 1972.
- Martin, F. *Introduction to audiology*. Englewood Cliffs: Prentice-Hall, 1975.
- Martin, F. *Pediatric audiology*. Englewood Cliffs: Prentice-Hall, 1978.
- Northern, J. *Hearing disorders*. Boston: Little, Brown and Company, 1976.
- Northern, J., & Downs, M. P. *Hearing in children*. Baltimore: Williams and Wilkins, 1974.
- Ross, M., Giolas, T. G. *TEJAS*, 1978, 1, 19-22.

Medical Evaluation: An Overview

Roger D. Freeman

Medical information may accompany the hearing impaired-child's records; however, it may be necessary to inquire of the family physician or pediatrician concerning specific medical problems of a child. There are two general principles to keep in mind with regard to medical evaluation: (a) Hearing-impaired children have a substantial likelihood of having other disabilities, recognized or unrecognized (Budden, et al., 1974; Jensema & Mullins, 1974); and (b) not all medically-labelled disabilities have educational significance. With these points in view, the following sections set out the kinds of information that may be necessary.

- **Age at onset of hearing loss.** Did the loss occur before the acquisition of useful speech? If so, there is likely to be much more impairment in speech intelligibility, especially if the degree of loss is severe or profound.

- **Audiologic findings.** These may be included in a medical summary or there may be a separate report from the audiologist. School health personnel should have a copy of the most recent audiogram. What kind of loss is it? Conductive losses may be surgically correctable and do not affect the fidelity of sound transmission; sensorineural losses are not surgically correctable, and amplification may not completely compensate for the loss. What is the severity of the loss and at what frequencies? This will be shown on the audiogram, and it is one indicator of the extent to which amplification and auditory training may be useful. It may also suggest that other means of communication may be necessary. How helpful are hearing aids to the student? Again, some hints may be obtained from the audiogram.

- **Associated additional disability possibilities.**

1. **Visual.** Has vision been tested recently? In British studies of ordinary children whose vision was normal at age seven, 18% had significant problems at age 16 (Tibbenham, et al., 1978). Hearing-impaired children depend even more upon their vision; they must have the best possible correction of visual deficits and may need special attention to their class seating (Pollard & Neumaier, 1974).

2. **Cardiac.** Is there any restriction of physical activity indicated?

3. **Motor.** Cerebral palsy of mild to severe degree is not uncommon. Architectural barriers may be a problem. Communication by means of speech or sign language may be difficult, as may writing, and a language board could be required. Medical labels are insufficient (by themselves) to indicate how much the student can or cannot do.

4. **Mental retardation.** In what areas is the student limited? Is some specific medical syndrome, perhaps with other impairments, involved?

5. **Convulsive disorder (epilepsy).** What kind of seizures does the student have? Teachers are particularly concerned about convulsions but there may only be occasional staring spells or seizures during the night. Under what kind of control have they been? Will drugs, especially if the student is taking several, impair attention in class? Must they be given during school hours? Are there any special precautions or limitations? Swimming may need to be more closely supervised;

lack of physical exercise is often worse than the reverse, in terms of frequency of seizures.

6. **Other types of brain damage.** A central language disorder affecting processing of information may be difficult to identify but could be one factor in producing learning problems. Sometimes this will be indicated by psychological testing or through the results of special tests, but those conducting such tests must be experienced in their application to hearing-impaired children. It is essential to remember that the presence of some form of brain damage, such as cerebral palsy, does not in itself explain all the learning or behavior problems a student may have; there needs to be independent evidence of impairment in the particular brain function in question. The use of the terms "minimal brain dysfunction," "minimal brain damage," "learning disabilities," and others is so variable from professional to professional and from center to center that it cannot be relied upon to provide the educator with useful information. One must determine what the specific abilities and disabilities are under specific types of circumstances before making educational decisions.

- **Chronic medical illnesses.** Any illness that may impair a child's attendance, attention, and level of participation in school activities should be known to the school (Bryan, et al., 1978). Two examples of such illnesses are kidney disease and diabetes. Explanations should be obtained from the responsible physician as to severity, implications, and treatment relevant to the school setting.

"Most studies show that 20-40% of hearing-impaired children and youth, as a group, have additional disabilities."

- **Psychiatric disorder.** Does the student have known problems of behavior or mood? If so, how does this affect school learning and social functioning? Is there some way in which the school can assist in management, such as by making observations useful to the mental health professional? Are any psychotropic or other drugs being used? What are their effects? Behavior disorder is usually acknowledged to be somewhat more frequent in hearing impaired children. It is important to recognize that few psychiatrists—or other mental health professionals—have any training or experience with the hearing impaired. Their pronouncements must therefore be taken with greater than usual caution (Freeman, 1977).

- **Drugs given for other reasons.** Medication may be used for allergies at certain times of the year, and these may influence alertness.

In all cases where there is a significant medical treatment or problem, the name and telephone number of the supervising physician should be available.

Most studies show that 20-40% of hearing impaired children and youth, as a group, have additional disabilities. The distribution of these problems depends to some extent upon the cause of the hearing loss. In rubella children, almost half have cardiac defects; about one-third have visual problems; and about one-fifth are of below average intelligence. Meningitis as a cause of deafness is associated with a high rate of mental retardation. One such survey showed that one-third of all deaf children had significant visual problems (Pollard & Neumaier, 1974).

School administrators can make wiser decisions about placement if



they consider not only what medical problems the hearing impaired child may have, but also the practical consequences or implications of these problems upon the student in the educational setting.

References

- Bryan, E. et al. Medical considerations for multiple-handicapped children in the public schools. *Journal of School Health*, February 1978, pp. 84-89.
- Budden, S. S. et al. Deafness in infants and preschool children: An analysis of etiology and associated handicaps. *American Annals of the Deaf*, 1974, 119, 387-395.
- Freeman, R. D. Psychiatric aspects of sensory disorders and intervention. In Graham, P. (Ed.), *Epidemiological approaches in child psychiatry*. New York: Academic Press, 1977.
- Jensema, C., & Mullins, J. Onset, cause, and additional handicaps in hearing impaired children. *American Annals of the Deaf*, 1974, 119, 701-705.
- Pollard, G., & Neumaier, R. Vision characteristics of deaf students. *American Annals of the Deaf*, 1974, 119, 740-745.
- Tibbenham, A. D. et al. Vision screening in children tested at 7, 11, and 16 years. *British Medical Journal*, 1978, 1, 1312-1314.

Psychosocial Needs of Deaf College Students from a Faculty/Staff Perspective

Jenny L. Covill-Servo
Wayne M. Garrison

Personal/Social Needs Assessment

Traditionally educational curricula have emphasized the development of cognitive and technical competencies. However, as more and more hearing impaired students have been absorbed into the hearing world of work, it has become increasingly important to examine the development of students' personal/social competencies and the influence of psychosocial variables on the future job satisfaction and competency of the deaf worker.

In keeping with these concerns and specifically, in preparation for a course entitled "Applied Human Relations," a needs assessment was conducted at the National Technical Institute for the Deaf (NTID) during the winter of 1978. The purpose of the survey was to determine the perceived personal and social problems confronting most NTID students. It was assumed that the faculty and staff of NTID would serve as a rich resource in providing useful information regarding the personal/social competencies that needed to be addressed either in the "Applied Human Relations" course or in the context of other courses taught at NTID.

Needs Assessment Questionnaire

The Needs Assessment Questionnaire contained five compound questions; four of these were open-ended, and one required respondents to rate the severity of various problems. Question #1 asked the respondents to provide basic information about themselves, including the number of years experience working with the deaf as well as the nature of their

positions. Question #2 requested the respondents to indicate what they perceived as the *major* personal and social problems confronting *most* NTID students. Question #3 asked that they indicate which topics they would address in a course devoted to the teaching of personal/social competencies. Question #4 contained 16 items in the form of a list of general problems. The respondents were asked to rate (on a scale ranging from 1 to 5) the degree of emphasis which they personally would give each problem area in a course of study. Question #5 was open-ended and asked the respondents to indicate which instructional techniques they had found to be most useful in teaching deaf adolescents. Space was provided at the end of the questionnaire for respondents' comments. A complete copy of the survey form is contained in Appendix A.

Results

Return Rate

A total of 89 questionnaires (approximately 15% of those distributed) were completed and returned for analysis. There are several reasons why the response rate may have been so low:

1. One department (containing about 40 people) declined to participate, stating that their Code of Ethics prevented them from revealing any information on specific individuals. This was a point of confusion, since the Needs Assessment Questionnaire did not request information on individuals but merely general impressions concerning problems that were shared by many deaf adolescents. Other

individuals also may have been confused about the level of information desired.

2. Some people objected to the structure of Question #4, arguing that the items were phrased in terms of stereotypes with which they disagreed.
3. Office staff may have had the false impression that their comments were not desired.
4. Others may have been too preoccupied to complete the questionnaire or may have felt the information would not be put to any useful purpose.

Question #1: Background Information on the Respondents

Of the 89 questionnaires returned, 47 were from teachers, 18 from administrators, 9 from researchers, 4 from counselors, and 11 from persons in miscellaneous positions. The highest rate of return was from the teaching group, probably reflecting the fact that there are more teachers than other kinds of professionals at NTID.

Information provided in Question #1 indicated that over 50% of those persons who returned the questionnaire had worked with the deaf for four years or less. These data probably reflect the relatively recent introduction of many professionals to the field of deafness, rather than the lack of interest of persons who have worked with the deaf for longer periods of time.

Question #2: Major Personal/Social Problems Confronting Most NTID Students

Question #2 was answered by 88% of the respondents. Although there was wide variability in the perceived problems, the general consensus

"Less frequently cited problems included . . . failure to consider the consequences of one's behavior . . ."

Table 1
Perceived Personal and Social Problems
Confronting MOST NTID Students*

Problems cited	Frequency
1. Lack of responsibility	23
2. Immaturity	13
3. Lack of information and experience	13
4. Lack of confidence	12
5. Poor decision-making skills	12
6. Poor interpersonal communication	9
7. Lack of sensitivity to others	7
8. Lack of social etiquette	7
9. Poor self-concept	7
10. Lack of self-awareness	5
11. Dependency	3
12. Inflexibility	3
13. Poor conflict resolution skills	2
14. Unrealistic goal-setting	2
15. Lack of risk-taking behavior	2
16. Failure to consider the consequences of behavior	2
17. Poor family relations	1
18. Malleability	1

*This was an open-ended question to which more than one answer was frequently provided.

Table 2
Suggested Course Content

Topics	Frequency
1. Accepting responsibility	15
2. Decision-making	11
3. Interpersonal communication	11
4. Self-awareness	10
5. Survival skills	10
6. Conflict resolution	6
7. Interpersonal relations	6
8. Social etiquette	5
9. Independence	4
10. Friendship formation	4
11. Appropriate on-the-job behavior	3
12. Time management	2
13. Risk-taking behavior	1

seemed to be that the greatest problems facing deaf adolescents are (a) lack of responsibility, (b) lack of information and experience (immaturity), (c) lack of confidence, and (d) poor decision-making skills. Less frequently cited problems included (a) failure to consider the consequences of one's behavior, (b) lack of risk-taking behavior, (c) inflexibility, and (d) dependency. Table 1 provides a complete listing of problems cited by the respondents, together with the frequency with which they were mentioned.

Question #3: Recommended Topics to Address in a Course Designed to Enhance Personal/Social Competencies

Question #3 was answered by approximately 80% of the respondents. The frequencies with which various topics were suggested are shown in Table 2. The most frequently mentioned topics included (a) accepting responsibility, (b) decision-making, (c) interpersonal communication, (d) self-awareness, and (e) survival skills.

Question #4: Problems Confronting Deaf Adolescents

Question #4 was the only rating question included in the Needs Assessment and, perhaps, the most difficult one for people to answer. As the data in Table 3 indicate, when asked to rate problem areas in order of priority those given the highest mean ratings were (a) decision-making ($\bar{x}=4.37$), (b) discriminating those problems which were unique to deaf individuals ($\bar{x}=4.13$), (c) conflict resolution ($\bar{x}=4.09$), (d) survival skills ($\bar{x}=3.97$), and (e) obtaining reliable information ($\bar{x}=3.94$). Areas that were consid-

"The techniques that the respondents found to be most useful were those which required student participation . . ."

Table 3
Mean Ratings on Items Contained in Question 4

Statement themes	Mean
E. Decision-making	4.37
L. Discriminating problems	4.13
G. Conflict resolution	4.09
K. Survival skills	3.97
F. Obtaining reliable information	3.94
B. Interacting with hearing people	3.64
N. Articulating feelings	3.61
D. Kinds of relationships	3.54
I. Impact of behavior	3.46
O. Self-awareness	3.33
M. Self-image	3.24
C. Self-assertion	3.08
A. Self-defeating behavior	3.06
P. Sex education	2.74
J. Future anxiety	2.42
H. Invalid assumptions	2.16

ered to be of lesser importance were (a) sex education ($\bar{x}=2.74$), (b) future anxiety ($\bar{x}=2.42$), and (c) invalid assumptions concerning the reactions and conversations of hearing persons in close proximity ($\bar{x}=2.16$).

Question #5: Most Useful Instructional Techniques

Question #5 was answered by 51% of the respondents, which is not surprising since 53% of the total pool of respondents were teachers. The techniques that the respondents found to be most useful were those which required student participation, visual aids, and role-playing. The lack of consensus regarding the effectiveness of various techniques probably reflects the limited number of techniques utilized by instructors rather than the comparative efficiency of the techniques.

Additional Comments

Very few people provided any ad-

ditional comments (19%). However, most of those who did, expressed an interest in discussing these issues in greater detail and/or being of future assistance.

Conclusions

The data obtained in response to Questions 2, 3, and 4 indicate that most respondents felt the greatest personal/social problems confronting NTID students were lack of responsibility and poor decision-making skills. Rather than constructing a specific course to deal with these issues, it is suggested that many courses presently offered at NTID could be structured to require students to accept more responsibility and to make more decisions on issues affecting their lives. Many respondents felt that NTID students still had difficulty with interpersonal communication and survival skills; this

suggests that courses presently offered in these areas may need to be offered more frequently or recommended more often by the counseling staff. Some topics which seem appropriate to address in a course on "Applied Human Relations" (which do not seem to be presently addressed by other courses) include self-awareness, friendship formation, and conflict resolution techniques.

As a consequence of the study, the decision was made to develop a two-course sequence on Human Relations. Work on *Human Relations I: A Self-Awareness Course for Hearing Impaired Adolescents*, has proceeded beyond the initial phases of development. This includes the production of a student manual to accompany the course. The course *Human Relations II* will be developed in the near future and will include such issues as conflict resolution, friendship formation, and various other techniques to improve relationships with others.

Appendix A
Needs Assessment Questionnaire

**Needs Assessment: Psychosocial Development
of NTID Students**

We are revising our course entitled "Applied Human Relations" to more adequately meet the personal/social needs of NTID students. Many of you are (or have been) working directly with the students and have a good knowledge of the strengths and weaknesses which they possess in the personal/social domain. Your help in defining the needs that could be addressed in our course would be most valuable. Please take a few minutes to respond to this questionnaire and return it to J. Covill-Servo, Department of Research and Development, Building 50, Room B-103.

1. Name (optional) _____

Years of professional working experience with deaf individuals _____

Nature of your position (check the ONE which BEST applies):

Teaching _____
Counseling _____
Research _____
Administration _____
Interpreting _____
Other (specify) _____

2. From your own experience, what are the *major* personal and social problems confronting *most* NTID students?
3. If you were offering a course to enhance the personal/social competencies of NTID students, what 2-3 key topics would you address and why (briefly)?
4. In conversations with several faculty and staff at NTID, a number of general problems have emerged which could be addressed in our course. Please read the statements below and indicate for each the degree to which you personally would give emphasis in a course of instruction using the following scale:

1 2 3 4 5
└─────────┴────────┴────────┴────────┴────────┘
LOWEST low ? high HIGHEST
EMPHASIS SCALE

In determining the extent of emphasis which you would place on each of the areas below, consider the target population to be MOST NTID STUDENTS. Please read the ENTIRE list of statements BEFORE responding to specific items.

- a. Many students are afraid to try new things because they think that they are not capable of succeeding. They are *self-defeating*. _____
- b. Students need to feel more comfortable and be more skilled in *interacting with hearing people*. _____
- c. The students need to *assert themselves more frequently*. _____
- d. The students need more *information about dating, marriage, friendship formation, and kinds of relationships* in general. _____
- e. The students are not accustomed to *making decisions for themselves*. _____
- f. Students do not know where to go for *reliable sources of information* when they encounter problems. _____
- g. The students need to learn how to *resolve conflicts* and cope effectively with certain dissatisfactions which arise in their daily activities. _____
- h. In general, many students assume that when hearing people in their proximity are talking amongst themselves, they are saying derogatory things about the deaf students. Their *assumptions* may be invalid. _____
- i. The students need to be more aware of the *impact of certain aspects of their behavior* on hearing people (e.g., loud music, yelling, blocking hallways). _____
- j. The students express *anxiety* about the future. _____
- k. The students need to learn *survival skills* (e.g., how to open a bank account, pay bills, maintain a budget). _____
- l. The students can't discriminate which *problems are related uniquely to their deafness* and which are problems that most people face. _____
- m. Many students have a *poor self-image*. _____

n. Students don't know how to *articulate their feelings*. _____

o. Many students show signs of a *lack of self-awareness*. _____

p. The students would benefit greatly from increased
attention to *sex education*. _____

5. (IF APPLICABLE) What instructional techniques have you personally found most
useful in teaching your own content area(s)?

6. Additional Comments:

THANK YOU FOR YOUR HELP!

Assessing Labor Cost of Objectives

Siegfried Engelmann
Alex Granzin

If the assessment of the child's skills is to lead to instructional remedies of some sort, the assessment should imply the "cost" of the intervention. How much time or effort will it probably take? Can it be achieved easily through a few demonstrations or can we inevitably expect to devote hundreds and hundreds of trials before achieving even acceptable approximations of particular skills?

If the teacher is to prioritize teaching and make intelligent decisions about when to schedule various skills, these questions must be answered. Teachers are often reasonably knowledgeable about the teaching of academic skills, such as reading and arithmetic. They know, for example, that it would take years of pretty good instruction to bring the perfectly naive learner (deaf or otherwise) to a reasonable understanding of equations or to achieve good reading performance. Even though the teacher may not have a precise idea of the number of hours or trials required to induce each particular skill, the teacher knows that it would not be wise to schedule reading this week and after achieving it turn to arithmetic and allow a week for that.

When we move to the more basic areas of speech production and language concepts of the profoundly deaf child, we generally observe far less understanding of what it will take to achieve goals and changes in behavior. The approach is sometimes glib, providing a few trials for something that would require hundreds, exposing rather than actually teaching, and apparently lacking priori-

ties. Children often practice what they can do but receive very little help in changing their behavior.

The purpose of this article is to outline procedures for assessing the potential labor cost of changing speech and language behavior of profoundly deaf children. The procedures are not expressed as test items but as ways of inferring the nature of the skill from the performance of the child. This assessment, in other words, takes place within the context of the instructional situation, not in the formal testing situation. It may be thought of as a cousin to or an extension of the learning potential procedures that first teach the learner something and then evaluate the performance. The labor-cost assessment techniques that we will examine require information about the learner's performance on different tasks. The learner's performance suggests whether the labor cost is very high or low.

Patterns of Responses

For this assessment, we have to begin with some ideas that may seem a little radical. When we teach language-related skills, we are teaching concepts. Concepts are nothing more than qualitative samenesses. The ultimate goal of instruction in cognitively related areas, therefore, is to teach the learner which samenesses to attend to, how to label them, and which conventional responses are designed to show them. The game of instruction is basically nothing more than showing the learner samenesses and differences. How are all occurrences of the sound /d/ the same? How does the /d/ sound differ as a function of different sound contexts? What responses do we require from

the learner to demonstrate knowledge of /d/ or proficiency with /d/? How is the sound /d/ different from other sounds, particularly those that are similar in their response structure (such as /t/ which is simply an unvoiced /d/)? If concepts are nothing more than qualitative samenesses, the learner should be able to demonstrate knowledge of a concept by responding to various positive examples of the concept with one response and to negative examples with another response. The learner, in other words, should demonstrate a pattern of responses that corresponds precisely to the pattern of samenesses. When the sameness is present, the learner produces one response. When it is absent, the learner produces the other response.

This pattern is quickly induced for all concepts if the learner has no problem perceiving the samenesses and differences and has no problem producing the motor responses that we require for the task. The communication with the learner must be capable of demonstrating the samenesses and differences and doing so in an unambiguous way. If the communication is unambiguous, however, the pattern will be induced very quickly, within a minute or two. Later, the learner may forget the response but the pattern can be induced even faster now.

Here is an illustration: Let us say we are working with a child who does not know the meaning of the word *on* when it is used to describe the position of objects. Let us say that we provide a demonstration that presents the same objects (a ball and a table) for a series of examples, and the only difference from example to example is the position of the ball. We appropriately label four examples (on . . . not on . . .), and we

The major support for the research reported in this article came from the Bureau of the Handicapped, U.S. Office of Education; Grant #G007604355.

"After the initial communication, we provide the learner with a series of test examples . . ."

are careful to control for the order of the examples (so they do not alternate from "yes" to "no," and so they show a range of positive variation).

After the initial communication, we provide the learner with a series of test examples, still involving the ball and table. We present examples that had not been presented during our demonstration and initial practice. Let us say that we test the learner on 10 items. Performance on this test confirms that we have induced a pattern of responses that is based on the samenesses and differences of the examples we present. The learner's responses, in other words, provide a parallel to the samenesses of the examples. All test examples that have the property of showing the ball on the table are responded to one way while all test examples that do not have this sameness or quality are responded to another way. A summary of the learner's performance on the test items is found in Table 1.

When asked the question, "Is it on the table?" the learner responds one way to items 2, 3, 4, 7, 8, 10 and another way to items 1, 5, 6, 9. After an intervention of perhaps a minute, a pattern is induced. That we have induced a pattern and not a series of rote responses can easily be demonstrated by testing the learner on other examples involving the table and the ball. This time, however, we require a different response—touching all the examples that are *on* (instead of answering "yes") the table. We note that the same pattern is produced. For all positive examples, the learner produces the touching response; for negatives, no touching response. If the pattern persists across different responses, we have induced a pattern.

Table 1

Positive examples	2	3	4	7	8	10
Yes responses for examples	2	3	4	7	8	10
Negative examples	1			5	6	9
Negative responses for examples	1			5	6	9

In some areas of instruction, the child can be expected to perform in this manner—assuming that our communication is not faulty and does not suggest possible misrules. For instance, in teaching the learner the relationship between fractions that have a denominator of 100 and decimal notations, the presentation of perhaps 20 examples of the form below would induce the understanding of sameness and permit the learner to generalize to countless examples.

$$\begin{array}{r} 20 \\ 100 = .20 \\ 5 \\ 100 = .05 \\ 7.36 \\ 100 = .0736 \end{array}$$

Many instructional objectives are of this type. We show the pattern and perhaps review it on subsequent lessons, and the pattern is induced.

Inferences on Pattern Distortion

To make the pattern more difficult for the learner, we can operate in two different domains.

1. We can make the samenesses and differences less perceptible or less obvious. (The learner will have more trouble "seeing" what controls the pattern.)

2. We can specify more difficult responses. (If the learner can not

produce an acceptable response, the learner fails the item even if the learner apparently perceives the pattern.)

By combining the two ways of increasing difficulty (making the samenesses more difficult to perceive and requiring difficult responses), we create the most difficult learning.

Each type of difficulty has a very distinctive distortion. A test on a series of examples gives a pretty good idea whether the concept is "simple," whether the learner has a problem of perceiving samenesses, whether the learner has response-production problems, or whether the difficulty involves both perception and response problems.

Let us say that we present a very noisy presentation of *on* to the learner. We show the same set of examples that we showed before, but we interrupt the presentation with observations about other things. We talk about other concepts, (the ball, the table, etc.), and we generally make it less obvious about the sameness observed by various examples that are labelled "*on*." In this situation, we induce the pattern quickly; however, the pattern is wrong, which means that the learner's pattern of responses on the test items does not correspond to samenesses and differences of the items (see Table 2). The learner's pattern of producing positive responses and negative responses is quite different from the pattern of examples. The communi-

"If the learner has motor-response problems, we know that we must shape the response . . ."

Table 2

Positives	1	3	4	6	9	10
Yes responses	1	2	3	5	7	8
Negatives			2	5	7	8
No responses			4	5	6	10

cation failed not because of the learner's inability to produce the responses, but because the learner apparently could not perceive the samenesses and differences that signal positive and negative examples. This distortion shows discrimination difficulties.

The pattern for motor difficulties is quite different. The learner produces responses that are not called for by the task. (The learner may be required to say, "on" or "not on.") The actual responses the learner produces for *on* and *not on*, however, are "uh" and "dah uh." Indeed, we can tell whether the learner is trying to say one response or the other. However, our task objective may be quite specific about the kind of responses the learner is to produce. The responses above are approximations of responses rather than responses specified by the objective signal motor-response problems. The extent to which the motor-response problem is independent of the sameness-perception problem can be ascertained by looking at the pattern of responses. If the learner's two responses vary as the examples vary, the learner perceives the sameness but is incapable of responding with the specified responses.

A final bit of information permits us to complete our assessment procedure. Both problems that involve sameness perception and motor-response problems are not corrected quickly. Instead of dealing with a cursory presentation of 20 examples

or trials, we may be looking at several hundred trials.

Our identification of the problem suggests the direction and nature of intervention. If the learner has motor-response problems, we know that we must shape the response, accepting approximations and changing our criterion for reinforcing responses until the learner consistently produces acceptable responses. This remedy is implied by a pattern of responses that are approximations but that parallel features of the examples.

If the responses are formed acceptably but the pattern is wrong, a totally different approach must be used. We must somehow make the samenesses and differences more perceptible or obvious to the learner.

If the learner's pattern of responses discloses both an inappropriate pattern and unacceptable responses, we are faced with a learner who does not know exactly when a response is called for or how to produce the response. To show *when*, we have to make the samenesses and differences of the examples more perceptible. To show *how*, we must shape the response. Then we must combine when and how.

Characteristics of Difficult Learning

The type of remedies are derived fairly clearly from the nature of the problem. If the learner's pattern of responses suggests sameness-percep-

tion problems or motor-response problems, we are provided with important assessment information. No matter how carefully we proceed, however, the intervention that we provide will encounter problems to at least some degree.

1. The learner will initially show no progress. Even with the careful control of reinforcers and a careful presentation that shows what to attend to, the learner may not catch on to what we are trying to convey.

2. After catching on, the learner will demonstrate frequent regressions and very erratic performance. For a few moments the learner may show signs of responding to the samenesses of the examples. With no apparent reason the learner may then lapse into a "guessing" pattern. Similarly, the learner may produce a response that is nearly acceptable on one trial. The probability of the learner producing a response as good on the subsequent trial, even if we present it immediately after reinforcing the learner, is very low.

3. Changes in the context in which the discrimination or response occurs disrupt performance on skills that had been "firm" in more simplified contexts. If we introduce a new word for the learner to say, the learner may have trouble not merely with the new word but with some words that had been produced quite acceptably before the introduction of the new word.

4. The learning of skills that are of the same type that we initially teach will proceed faster and faster. The initial skills may require hundreds and hundreds of trials. The learner may subsequently learn a new skill in only one or two trials. However, this outcome is possible only if we first bring the learner to an acceptable criterion of performance

"Most instruction in speech production and speech perception was provided through a vocoder, a device that transforms sound into patterns of tactual vibration."

on the earlier skills. There is no magic to transfer. Transfer happens only when transfer is possible, which means the examples are the same in some way. Not only must transfer be possible before we can see an accelerated rate on later skills, but also the learner must be taught the initial skills.

Case History

When we apply the assessment of potential labor implied for achieving objectives to deaf children (particularly in the areas of speech and language skills), we see from their pattern of responses (or nonresponses) that they have serious discrimination problems which are very frequently compounded by motor problems. These assessments provide a warning that the problems noted above will occur during instruction and that we will achieve our goals of thorough learning, generalization, and facility with the various skills only if we accept the extremely intensive task of providing a skill foundation.

We can illustrate the various problems through a case history of a deaf child whom we taught speech and language skills using a vocoder. The use of the vocoder enabled us to keep relatively precise control over the input the youngster received. The child received information from the vocoder that was not available without it, furthermore, the information provided by the vocoder was adequate to show samenesses and differences.

The child was Ann, a 13-year-old profoundly deaf child who had been in a state institution for nine years and who had been labelled "multiply handicapped" (for no reason that

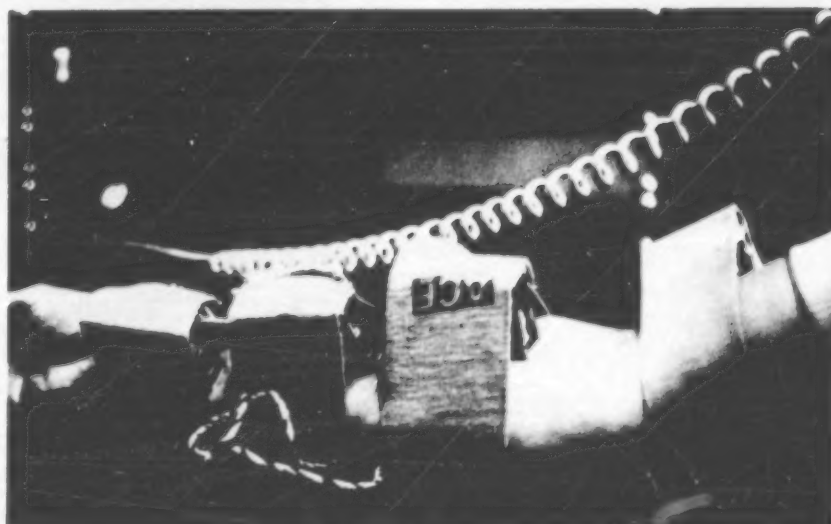
was apparent to us). She was quite naive and enormously deficient in social as well as instructionally related skills.

Most instruction in speech production and speech perception was provided through a vocoder, a device that transforms sound into patterns of tactual vibration. The device consists of microphones (one for the teacher and one for Ann), the vocoder, and 24 vibrators, each of which respond to energy in a specific frequency range. The instructor speaks into the microphone. The message goes through the vocoder where the speech is divided into 24 channels (spanning 90HZ to 1200-HZ). The pattern of energy goes to leads that are in contact with Ann's thighs, where vibrators are activated. If there is energy in a particular channel, the vibrator for that channel vibrates. When Ann speaks, the same process occurs. The voice goes through the vocoder and activates the various vibrators. Ann is there-

fore provided with a basis for comparing what she says with what the instructor says. (Match the vibratory pattern and you match the sounds.) The tactual display is laid out so that the left distal vibrator recorded the lower frequency sounds while the right proximal vibrator processed highest frequency sounds (/s/).

During most of speech production instruction, the instructor faced Ann. For the speech perception, however, the instructor either (a) held a clipboard in front of his face, or (b) sat behind the subject. In either case, a minimum of facial or other spurious cues were available. For speech perception training, the instructor said words at a normal speaking rate. Ann simply identified each word.

The speech production program used with Ann involved teaching a perfectly phonetic orthography and using it to work on the production of sounds in isolation, words in isola-



"The tasks . . . did not provide Ann with information that her strategies for 'learning' were inappropriate."

tion, patterns of repeated "words," sentences, and patterns of repeated sentences.

The advantages of orthography were:

1. The orthography permitted us to provide prompts for producing specific sounds. It is sometimes difficult to communicate which sound within a word is incorrect. Without this prompt, the learner may not perceive the difference between the wrong way and the right way even when you stress and exaggerate it (thereby running the risk of distorting it so badly that the learner will never recognize it). In situations like this, it is very convenient to present a pronunciation guide for the word or utterance.

2. The writing code permitted us to work in two directions at the same time. We could establish some responses fairly perfectly in a simplified context—responses to individual symbols. We could then change the context for these sounds and start to prompt generalizations involving responses that are produced quite acceptably. This is an atomic approach. At the same time, we worked on a more macroscopic level, reinforcing approximations of utterances. Here we were working on the attempts to imitate a variety of spoken utterances.

In many respects, the instructional context for teaching Ann was ideal for observing and trying to correct specific problems. The vocoder provided us with the means for showing exactly how to produce responses. Even with these controls and with good instruction with reference to the pacing of the tasks, specific corrections, use of reinforcement, and skill integration and application, Ann exhibited the full range of problems predicted by her initial performance.

She experienced both perception problems and motor response problems.

Initial Inability to Perform

When Ann first began instruction, she had a host of apparent misrules about instruction, speech communication, and learning. She did not know signs or fingerspelling. She had no oral speech. Her behavior was characterized by smiles and attempts to imitate mouth movements. If her instructor attempted to give her instructions, she would try to imitate facial movements, usually uttering no sounds. For her, learning was apparently a game of trying to figure out what the instructor wanted her to do. She would not attend to the features of examples presented to her. Instead, she would try to read the instructor's face or would try some sort of guessing strategy.

The initial instruction was designed to acclimate her to the instructor, the vocoder, and the format of instruction. The tasks were not particularly demanding and therefore did not provide Ann with information that her strategies for "learning" were inappropriate. From a purely logical standpoint, the most effective initial instruction would have been that which provided her with instant and dramatic information about the inappropriateness of her strategies. The problem with that approach is that it would have been very punishing because virtually everything Ann did would have led to negative consequences (or lack of reinforcement).

After the acclimation phase of two or three sessions, more concerted instructional sequences and tasks were introduced. Ann had been presented

with oral words such as *mother, elephant, shoe, pop*. She was provided with pictures showing each word. When the instructor (who stood behind her) said a word, Ann pointed to one of the pictures. Her pattern of performance did not even closely approximate the samenesses and differences in words. She was not attending to samenesses in the tactual words. Stated differently, the perception tasks were very difficult for her. After approximately 500 trials with the set of words above, the verdict was clear that Ann was attending to features other than samenesses in vibratory patterns.

The next step involved presenting examples that are more highly discriminable than the words above. This activity did not replace the work with the words but was supplemental to it. The instructor said two sounds—the /m/ sound and the /s/ sound. These sounds were represented as the very distinct vibratory patterns, with the /m/ sound on the distal left leg and the /s/ sound on the proximal right leg. The instructor would say a sound and Ann was to point to the appropriate leg. This task was apparently simple. Ann received about 400 practice trials with /m/ and /s/ presented in an unpredictable order. After each mistake, Ann was corrected. At the end of 400 practice trials, her performance was still random, even though every instructional and equipment precaution under our control was taken. Next, the instructor removed all vibrators from the right leg (the bank that transmitted the /s/ sound). Now the task was a simple discrimination task of pointing when the /m/ sound was presented. On the first trial, the instructor said /m/ and Ann pointed to the vibrator. She received reinforcement. Again the

"... if a response had just been produced, there should be substantial memory of how it was produced ..."

instructor said /m/, and Ann pointed to her right leg. It should be noted that during earlier training, items had not been presented in alternating order. When we compare Ann to other learners, we may "tiak tiak" over her incredible naivete. Comparisons of this sort are not relevant. We did our best to make the perception of the samenesses as obvious as possible. Ann's performance, however, suggested that they were anything but obvious to her. The basic remedy is practice, feedback, and patience, with the understanding that in time, Ann would learn the system. After about 200 trials with vibrators on only one leg, she performed accurately, pointing only when the sound was produced. For the first time, Ann demonstrated a response pattern that corresponded precisely to the pattern of examples.

Regressions and Erratic Performance

The initial learnings involve experimentation. The learner must somehow learn rules for producing a particular response and for attending to a certain type of feature. Before formulating appropriate rules, the learner may have to discard many abortive rules. The most dangerous abortive rules are those that permit the learner to perform on some easy examples, but that do not work for a broader range of applications.

Ann had predictable problems producing different sounds. She had less trouble with those sounds that are easily produced if the mouth assumes a certain position. She had no trouble with the /oo/ sound. However, Ann had serious motor problems with some sounds for which

there was a clear "marker," such as the /ē/ sound. (When this sound is produced, a shelf is formed just below the bottom jaw, and it should be about the only sound for which this shelf forms.) Even though Ann's instructor faithfully used this prompt by holding Ann's hand against his neck as he said the sound and then placing her hand against her own neck as she said it, Ann could not consistently produce an acceptable /ē/ sound after approximately 1,000 trials. By then, she could perform acceptably on perhaps 2 of 3 trials. However, the context in which the task occurred made a great difference on her performance. If the instructor modelled two or three examples and then presented the task to her: "Say eeeee," the probability of a correct response was much greater than it was if the instructor did not precede the task with prompted trials. Without the prompted trials, the probability was about 50% that Ann would produce the sound /u/ with a smile and squinted eyes, a facial prompt that had often helped her produce the /ē/ sound.

Even after Ann had produced two or three consecutive /ē/ sounds—with clear acknowledgment from the instructor that she was doing a great job—the probability of the next /ē/ being appropriate was only about 2 of 3. The probability of her producing an appropriate /ē/ after a pause of 30 seconds, during which the instructor pretended to be writing information on his chart, was about 50%. Sometimes, the responses that she produced after such a pause would suggest that she had never produced the sound before in her life.

Ann behaved in a perfectly lawful manner. The law involved here is that the probability of a correct response is increased if that response

can be produced within the easiest response context. The easiest response context is the one that shows the correct response or one in which the learner has just produced the correct response. Because if a response has just been produced, there should be substantial memory of how it was produced, providing the subject is not fatigued and previous productions were of some consistency. Even when the response context is made as easy as possible, however, we note that the learner's productions are very fragile. Regressions are the rule, not the exceptions. Erratic performance is the essential characteristic of the learning, not an anomaly. The learner is trying to gain control of the motor mechanism and to tie this mechanism to the prompts provided by the instructor. The process may fail because the learner cannot produce the response or because the learner forgets how to link the response with the signal provided by the instructor.

Disruption Caused by Changes in the Context of the Discrimination or Response

One of the most serious faults associated with using "behavioral objectives" is that they tend to atomize instruction and prevent the integration of skills in a reasonable manner. They do not take into account that when the context changes, you may teach something now at the expense of other things later. Let us say that the instructional objective is to teach the sound /d/. We work on the sound in isolation and bring the learner to an acceptable criterion of performance. We now change the context by requiring the learner to produce /d/ not merely in isolation

"During the early instruction Ann was requested to point to pictures that were presented."

but also within the context of words. The predictable outcome will be (a) deterioration of the /d/ response, and (b) deterioration of other responses that had been reasonably firm within the context of the words. When the learner produces the /d/ in isolation, the sound is quite different from the way it will occur in words, furthermore, the instruction fails to show the learner that the /d/ occurs in any contexts other than the isolated-sound context. Most isolated behavioral objectives lead to the same outcome that we can anticipate with /d/. We achieve the specified objective, however, the objective is perfectly unreasonable because it doesn't take into account the various contexts for the skills and, therefore, does not provide careful teaching of the skill within these contexts.

The same sort of disruption occurs if one teacher works with the child on speech and if the child practices particular responses only in one situation. The probability of the learner performing when the context is changed (the learner now required to respond to different teachers or in a different physical set up) is quite low but perfectly predictable. The initial instruction implies that the learner should learn what is needed to perform in the single-teacher situation, not in the contexts that are presented later. Conversely, if the extended contexts were expected of the learner, they should have been part of the early instruction.

When the progression of contexts is carefully controlled, however, we are faced with the same problem. If the learner has great difficulty with the initial contexts, we spend a lot of time working on it. We must establish the skill in some context, hopefully one that is fairly "easy." We

then change the context. This change in context predictably results in problems, not only with the skill that is being integrated into different contexts but also with other skills that had been firm.

Ann exhibited the full range of altered-context problems. If she was asked to say a familiar sound such as /oo/ when the vibrators were removed (but everything else about the context was the same), she did not perform consistently. Teaching her to perform skills taught during the lessons required daily reinforcement in which the instructor walked around the school with her immediately after the instruction and repeated those tasks that she could handle by attending to his face or tactual cues (touching his neck). Even after Ann had been in the program for three months, her performance in contexts outside the classroom was predictably lower than that in the classroom. If you met Ann in the hall during a different time of day and presented a task that was quite "visible," she would tend not to perform correctly. Instead she might smile and give some inappropriate mouth movements. Similarly, with motor response, if a familiar sound was placed in a new phonetic context, the probability of disruption was very great. After Ann's performance had been "acceptable" for a few weeks on the /ē/ sound in simple contexts, the sound /i/ was introduced. (The symbol for producing this sound was: oē.) Ann had already mastered the short o sound, which gave her no particular trouble in any context. To say the sound oē requires a simple combining of two familiar sounds. The instructor points to the sounds and says them: "oooooēēēē." The instructor repeats the sounding. The instructor then

gave Ann a turn (pointing to her). As he touched the letters she said, "oooooauuuuaa" with her tongue wavering uncertainly. Many trials later, Ann was still having trouble with the sound oē. Her problems, however, are not limited to this context. She has now begun to make errors with /ē/ in a variety of contexts, including contexts that had been "firm" for weeks.

The pattern of disruption for discriminations paralleled the response problems. Ann learned to identify the word *mother* within the discrimination context of the words *elephant* and *shoe*. When the word *man* was introduced and Ann was required to discriminate the words *man*, *mother*, *elephant*, and *shoe*, she could not initially discriminate between *man* and *mother*. However, as she worked on this discrimination, she began to confuse *mother* and formerly familiar words—*elephant* and *shoe*.

An issue related to the role of the context has to do with the relationship between discrimination and responses. During the early instruction Ann was required to point to pictures that were presented. During the same period, she was being taught to say the names of the objects. As soon as 10 items had been introduced in perception, the presentation changed so that Ann was required to say what the instructor said, not to point.

New words were introduced into Ann's perception vocabulary when she was able to perform on words already identified to a criterion of 80%. The words in the identification test were presented in an unpredictable order. If Ann achieved 80% correct on the test, a new word was introduced. If she did not achieve 80% correct, she continued to work on the existing set of words.

When the word set reached about

"Even though Ann had initially appeared to be an almost comic figure, she learned at an impressive rate."

40 words, Ann had serious problems. She could not achieve 80% correct on perception tests. Possibly she recognized the words spoken by the instructor but was simply unable to produce the responses for those words. What evidence was there for this diagnosis? She tended to miss the words that had been introduced most recently. In fact, her inability to achieve 80% correct was almost totally accounted for by the fact that she missed the most recently introduced 10 words with predictable regularity. Her average performance on the most recently introduced 10 words was about 20% (measured over a two-week period). Performance on the other words was usually above 80%.

To determine the extent to which she recognized the words but could not find the response that she wanted, we introduced a semi-specific prompt. Before the perception test, the instructor wrote these 10 words on a chalkboard and pointed to each as Ann read it. The instructor then presented the perception test without referring to the words on the board and without giving any clue at all about whether each word he said was written on the board. The test words were presented in random order—all 40 of them. If Ann recognized a word as one that was written on the board she was able to refer to the board and read it. The orthography gave her a guide for pronouncing the word. Therefore, the words on the board served as a response guide. For this guide to serve Ann however, she just had to recognize the word spoken by the instructor.

With the introduction of the semi-specific prompting procedure, Ann's performance on the most recently introduced 10 words jumped immediately from about 20% correct to

65%. Removal of the prompt resulted in her performance returning to about 20%. Note that these percentages are averages with high degrees of variability. On the first day that the prompting was introduced, Ann achieved about 90% correct with the most recent words. On some days, without the prompt, she would achieve close to 60%.

The point of this demonstration is that if the learner is dealing with a number of discriminations and these are new and fragile, a large part of the learner's problems may be associated with formulating links between those signals that come from the environment and specific responses that are to be produced. A signal calls for a response. Which response? How do you produce that response?

Learning Subsequent Discriminations

After a year of almost daily instruction, Ann behaved like a different learner. She learned new discriminations at a rate that was about six times the rate that she had learned them after being in the program for three months and probably over ten times as fast as she had learned them during the first three months. Not only was her rate increased, but the discriminations were more difficult because new perception words were added to a larger and larger set of words, a set which contained a larger and larger number of highly similar words. Following a long plateau when 40 words had been taught, Ann apparently worked out a system for identifying spoken words and producing the response she wanted. Her rate of new-word acquisition became about as fast as

any deaf child we had worked with.

Even though Ann had initially appeared to be an almost comic figure, she learned at an impressive rate. Her speech was far from adequate at the end of a year of instruction. However, her speech progress was as impressive as her perception. She had learned to say a variety of utterances so they were quite well phrased and quite understandable. She said, "I want a —," "That's a—," "Where is —?" "Do you have a —?" and similar observation patterns and manding patterns. She said short poems, such as "That's a fish; that's a dish; that's a fish on a dish." She could produce acceptable names for her 75 vocabulary words. She could produce most phonemes in isolation or in repeated patterns so they were quite acceptable. Her speech was not fluent, and she tended to regress to previous behaviors in those situations where she had extensively practiced previous behaviors. When communicating with her mother, she would tend to communicate by gestures, as she did when she entered the program.

Ann's behavior was very predictable. The amount of labor that was required to establish her early speech-perception and production behaviors was clearly signalled by her response patterns. She showed us, through these patterns, that the samenesses to which she was to attend were not obvious and that the responses she was to produce were not easy. She showed us that many of the easiest tasks we were prepared to present involved a combination of difficult discriminations and responses.

As we move from the careful setting in which Ann was taught to the typical classroom for the profoundly deaf, some things change. The prob-

"In brief, we may draw false conclusions about what can be done and about our role in doing it."

lems and predictions, however, remain the same. The response pattern of little Tony clearly suggests that speech responses are enormously difficult for him. His utterances are overlaid with nasality; his inflections are cacophonous; his breath control is almost lacking; his apparent understanding of what this lip-moving game is about is all but lacking.

We can play a dilettante role with Tony and his classmates by "sampling" speech activities, by going through the motions, and by using "liberal" standards for judging whether he has achieved a particular skill. We can collect isolated behavioral objectives and play with them, or we can approach the teaching with this understanding:

(1) Discriminations or responses may be difficult for the learner;

(2) The difficulty can be inferred from responses to a number of similar tasks;

(3) The assessment will tell whether the subsequent instruction will be labor intensive;

(4) The assessment will also suggest the direction of instruction, i.e., for motor problems, find the simplest response context and begin shaping. For discrimination problems, create a communication that makes samenesses and differences in the examples as obvious or perceptible as possible; and

(5) Finally, the assessment warns us about the types of problems that we will probably encounter: initial inability, frequent regressions, and disruptions in performance when the context changes.

Unless we heed these warnings, we will not provide intelligent instruction or intelligent diagnoses. We may assume that the learner has some sort of irreparable perceptual impairment; we may draw the conclusion

that the learner is incapable of learning; we may become irritated over the learner's lapses and problems with items that had been firm before we taught something new. In brief, we may draw false conclusions about what can be done and about our role in doing it. If we accept the task of teaching speech, let us do so with knowledge and realistic expectations, not with slogans or with instruction that does not even begin to provide the amount of repetition and practice demanded to induce motor skills and discriminations that are initially difficult for the learner. The first step is an assessment that tells us both what the learner's problems are and the amount of labor that will probably be involved in effecting a change.

Using Criterion-Referenced Tests with Handicapped Students

Lynne Anderson-Inman

Introduction

Educational history has witnessed a dramatic change in the assumptions educators make about learners. In the past, responsibility for acquiring the knowledge and skills necessary to perform successfully in the adult world was placed primarily on the student. Teachers were provided to direct the course of learning and pass out information but it was the students' responsibility to learn, regardless of the quality of instruction to which they were exposed. Students were assumed to differ widely in their ability to profit from instruction and tests were designed to sort out which students were which.

In the late 1950's and early '60's, psychologists and educators began to explore possible real-world applications for the principles of learning derived from laboratory research. Heavily influenced by B. F. Skinner's initial work with teaching machines and programmed instruction, educators and the general public have become increasingly aware of the power of good instruction. The blame for failing to learn has gradually been lifted from the student, as well-sequenced and systematically implemented instructional programs illuminate the potential for improving any student's academic behavior. With the realization that there is such a thing as "good" instruction came the need for teachers to evaluate the quality of their procedures. The advent of powerful instructional interventions and the subsequent desire on the part of the teachers to assess when their teaching had been effective established a need for alternate forms of educational measurement (Popham, 1975). Standardized achievement tests gave an overview of student improvement but

were too general to be of any use in demonstrating when instruction on specific goals had occurred successfully. Standardized diagnostic tests were useful in identifying areas of deficiency but were not particularly helpful in pinpointing specific skills requiring instruction.

The Advent of Criterion-Referenced Tests

The term "criterion-referenced" testing was first used by Glaser in 1963 when he drew attention to the need for an approach to measurement which would be sensitive to the effects of instruction. Glaser (1963) called for a test in which a student's score would provide "explicit information as to what the individual can and cannot do." Criterion-referenced tests were gradually accepted as more appropriate for making decisions about individual student performance than tests designed to obtain general information on the relative ability of a student as compared to other students. The latter type of testing, characteristic of most achievement measures currently used in education, has been labeled "norm-referenced" testing.

Comparison of Criterion-Referenced Tests with Norm-Referenced Tests

The most significant difference between criterion-referenced assessment and norm-referenced assessment is the manner in which students' scores on the test are interpreted. Scores obtained from norm-referenced instruments acquire their meaning by referring or comparing the individual's performance to that

of other individuals who have taken the same tests, i.e., a norm group. Student performance on norm-referenced tests is reported in a manner which helps to make this comparison clear. Percentile scores, for example, indicate the percentage of students above whom the student scored. Grade-equivalent scores communicate student performance level by referring to the average performance of previous test takers from a variety of different grades. Similar comparisons are implied by scores which are reported as age-equivalents or stanines.

In contrast, scores obtained from criterion-referenced tests acquire their meaning by comparing the individual student's performance on each skill with a previously established criterion or standard (Popham & Husek, 1979). How an individual student performs with respect to other students is considered irrelevant for criterion-referenced assessment. The criterion-referenced test is deliberately constructed to provide measurements for specific tasks or skills (Glaser & Nitko, 1971). The student's performance on each skill is compared with the criterion established as indicative of acceptable performance for that skill. If the student's performance is equal to or better than the pre-established criterion, the student is considered to have attained competence in or mastered that skill. Scores on criterion-referenced tests are most frequently reported as pass/fail for each skill identified, with "pass" indicating mastery to the predetermined standard of performance and "fail" indicating lack of mastery.

The manner in which test items are selected and retained is the second critical difference between norm-referenced and criterion-referenced

"Good test items are those which represent the content of instruction and are sensitive to effective instructional techniques."

tests and is directly related to their differences in purpose. Because norm-referenced tests are used primarily to evaluate student differences, i.e., identify how well students perform relative to each other; good test items are those which spread students out along some continuum (reading ability, for example). Test constructors try to select test items which maximize the variability in students' scores, the ideal being an item which 50% of the students answer correctly and 50% of the students answer incorrectly.

Criterion-referenced tests, on the other hand, have been developed to make evaluative decisions about instruction: (a) which skills a student has already acquired and therefore does not need to be taught, (b) which skills the student has not yet acquired and therefore needs to be taught, and (c) the effectiveness of specific instructional procedures on previously unlearned skills. Good test items are those which represent the content of instruction and are sensitive to effective instructional techniques. Test items which students answer incorrectly prior to instruction but answer correctly after instruction would be considered ideal.

There is a danger in using norm-referenced tests for making decisions more appropriately made using criterion-referenced instruments, and the danger rises from this difference in the way test items are selected. Teachers have occasionally tried to analyze individual student performance on norm-referenced achievement tests or standardized diagnostic tests item-by-item in order to gain information useful for instructional programming. This occurs most often when norm-referenced test developers have included a list-

ing of skills or objectives for each item as American Guidance Service, for example, has done with *Key Math* (1971). Test items selected because they add variability to the range of students' scores are not, however, accurate reflections of the skill domain they represent. In developing test items for norm-referenced assessment, the content and format of each item are juggled to achieve the desired variability in student responding. Such juggling is not usually conducive to the creation of test items which best represent the skill to be evaluated or to the inclusion of test items which will be sensitive to changes in student competence as a result of instruction.

Alternative Terms for Criterion-Referenced Testing

Before leaving this general overview of criterion-referenced measurement, a few words need to be shared concerning alternatives to the term "criterion-referenced." As with many popular movements in education, increased familiarity with a phenomenon fosters individualized interpretations by educators attempting to clarify the more subtle aspects of its nature. Such attempts at clarification have often led to the selection of alternative terminology in the belief that the new words will better reflect a particular interpretation and hence eliminate all previous confusion.

It is in this manner that terms such as "objectives-based" testing (Sullivan, 1971), "mastery" testing (Block, 1971, Crehan, 1974), "domain-referenced" testing (Denham, 1975), and "instructional-program based" testing (Becker & Engelmann, 1976) have surfaced. Al-

though differences among the concepts represented by these terms exist, fundamental to all above terms is an approach to measuring student performance which (a) specifies the behaviors to be measured, (b) establishes a minimum level of acceptable performance for each behavior, and (c) communicates information to the teacher which is useful for programming instruction and evaluating the effects of instruction. Although using the term "criterion-referenced" testing for this approach may be confusing to some and offensive to others, popular usage has established its right to be accepted.

Where Do Criterion-Referenced Tests Come From?

As the critical reader has probably already discerned, the term "criterion-referenced" refers more to a way of measuring and interpreting students' performance than it does to any specific test or type of test. In fact, the variety of tests claiming to qualify as "criterion-referenced" is astounding to the point of being suspicious. Developing quality criterion-referenced tests, however, is a difficult and time-consuming process. A growing source of criterion-referenced tests are the commercial publishers of other types of instruments for assessment and evaluation. Commercially developed criterion-referenced tests were slow in appearing, however. Test developers were wary of damaging their thriving market for norm-referenced tests and leery of the need to specify clearly what is being measured (Popham, 1975). The tests developed by commercial publishers are generally of two types: (a) those associated with massive instructional materials sys-

"Teacher-made criterion-referenced tests are most often developed by teachers of students with a wide range of academic problems . . ."

tems also published by the same company and (b) those designed to stand alone as diagnostic devices.

In response to the widespread lack of criterion-referenced tests and in concern over their frequent inappropriateness for some populations, numerous state and local educational agencies as well as some university affiliated projects have stepped into the test-making business. Although it is difficult for these part-time test developers to attain the rigor possible by the large commercial developers, locally prepared criterion-referenced tests have the advantage of matching the curricular preferences of a specific region or district. Criterion-referenced tests developed in this manner often reflect the orientation of the teachers who contribute to the tests' content and as such are usually similar in educational philosophy to the prevailing curriculum materials of the time and the geographic area. The tests designed by these groups or agencies have sometimes focused on skills felt to be essential for a specific student population or those with a particular type of handicapping condition. Locally developed criterion-referenced tests have also been growing in popularity as instruments for district-wide assessment (Olson, 1976) and statewide evaluation of minimal competencies for graduation from high school.

Criterion-referenced measurement techniques are also being used with increasing frequency in the development of teacher-made tests, and an excellent resource is available to those interested in designing criterion-referenced tests for classroom instruction (Gronlund, 1973). Teachers are attracted by the utility of criterion-referenced tests and the specificity of information they facilitate. Teacher-made criterion-refer-

enced tests are most often developed by teachers of students with a wide-range of academic problems, teachers who devote a portion of their curriculum to units of study on a specific theme, and teachers who feel the available criterion-referenced tests are incompatible with their approach to instruction or the needs of their students.

Using Criterion-Referenced Tests

The utility of criterion-referenced tests for teachers is derived from the fact that they are direct measures of student performance rather than indirect measures. For example, a student's ability or lack of ability to identify the main idea in a written paragraph does not need to be inferred from a general measure of "reading comprehension." Instead, the presence or absence of the skill can be tested directly using criterion-referenced assessment strategies. Direct measures of student performance are founded on the assumption that what one bothers to measure should be what one plans to teach or, of course, its corollary; what is taught should be what is measured. This approach to assessment necessitates careful specification of the behaviors tested or to be tested. It is exactly this precision, however, which results in information useful to teachers. Most criterion-referenced tests give the teacher pass/fail information for each behavior or skill tested. This information can then be used by the teacher to (a) assess the student's current level of functioning relative to the curriculum content, (b) pinpoint specific skills for immediate instruction, (c) monitor student progress through the curriculum, (d) probe for "mas-

tery" of skills subsequent to instruction, and (e) evaluate the effectiveness of the instructional procedures or program used.

A word of caution is needed at this point. It should be remembered that the purpose of criterion-referenced testing is to measure what is taught or what will be taught. Criterion-referenced tests are only useful to the extent that the behaviors measured correspond to the skills covered in the instructional program. In fact, it is only when testing and teaching cover the same well-specified skills that criterion-referenced tests can be of assistance to teachers for making instructional-based decisions.

Criterion-Referenced Tests and Mainstreaming

The special education teacher is no longer the sole professional responsible for programming and providing education to the handicapped students of this country. Federal legislation (P.L. 94-142) requires that every identified handicapped student have a written program of educational goals and objectives (Individualized Educational Program or IEP) developed by a multidisciplinary team, of which the special education teacher is only one member. In addition, the same legislation requires handicapped students to be educated with nonhandicapped students to the maximum extent possible, thus encouraging increased participation and responsibility on the part of regular education teachers. Although neither of these developments are totally new in concept or practice, both have magnified the importance of criterion-referenced tests in the education of handicapped students: the first because it neces-

"No student needs to be labeled or segregated in order to receive instruction at the appropriate level or appropriate pace."

sitates diagnostic information on each student's skill deficits, and the second because it requires new levels of data-based communication among teachers.

However, the extent to which and the conditions under which criterion-referenced tests prove useful in public school efforts to mainstream handicapped students varies considerably, largely as a function of school policy and teacher orientation to assessment. In schools or districts participating in large-scale efforts to individualize instruction, criterion-referenced tests are frequently administered to all students in a building on a regular basis for one or more tool subject areas. Tests adopted for such extensive assessment attempts are usually correlated to curriculum materials produced by the same publisher. Probably the best known example of the above are the tests and instructional materials developed by Glaser at the University of Pittsburgh's Learning Research and Development Center. Now published by Appleton/Century/Crofts, Glaser's *Individually Prescribed Instruction* (IPI) combines criterion-referenced assessment with self-instructional skill booklets to assist teachers in accommodating to individual student differences in math. A similar approach has been adopted by other publishers as well as by local educational agencies.

Use of criterion-referenced tests for the entire student body provides teachers with skill profiles for each student in a particular subject area, e.g., reading or math. This information is then used to program or group all students in the school for instruction in that subject, usually ignoring traditional grade designations. The advantages for handicapped students of a schoolwide system for as-

essment and individualizing instruction are probably obvious. The system is designed to accommodate learners at all levels. No student needs to be labeled or segregated in order to receive instruction at the appropriate level or appropriate pace. Handicapped students, tested using the same instruments, can therefore be integrated easily without special arrangements or modifications to the existing curriculum.

In the majority of schools, however, there is not a widespread use of criterion-referenced assessment for the entire student body. Criterion-referenced tests may be used primarily by the special education teacher or diagnostic specialist to gather in-depth information about students who are suspected of having problems with the academic curriculum. The tests most used for assessing current level of functioning and specific skill deficits are those designed to cover a broad range of behaviorally stated objectives for a variety of grade levels. These tests are usually developed to stand on their own as diagnostic instruments but frequently provide the teacher with prescriptive references to a number of commercially available materials, keyed by page to individual objectives.

The availability of criterion-referenced test data on handicapped students has considerable utility in a school's efforts to increase or maintain the extent to which students with learning and behavioral problems are educated with nonhandicapped students. Direct measures of the skills handicapped students have acquired facilitate an objective appraisal of their strengths and deficits, thus enhancing their prospects for appropriate placement within the mainstream of regular education. In-

dividualized Educational Programs (IEPs), developed with the aid of criterion-referenced test data and capitalizing on the specificity of information gathered, are far more likely to truly reflect a student's educational needs than IEPs constructed without such information. This in turn, leads to IEPs which are actually useful for instructional programming as well as functional for communicating with regular class teachers, ancillary personnel, and parents.

Criterion-referenced tests which provide the teacher with an individualized profile of skills are also useful in developing and implementing a cooperative system for monitoring student progress in multiple-class settings. As regular class teachers assume increasing responsibility for the education of handicapped students, efficient procedures need to be established for communication on a systematic basis between regular and special education personnel. Criterion-referenced test data, especially when closely correlated with the curricula of both special education and mainstream settings, can be invaluable in maintaining a well-focused cooperative effort towards meeting the instructional needs of handicapped students. Those tests which provide prescriptive information referenced to specific instructional materials are additionally useful by facilitating the search for supplementary materials on selected behaviors or skills requiring remediation. And finally, but not unimportantly, criterion-referenced tests can play a significant role in evaluating the success of selected instructional procedures, thus providing the feedback to regular and special education teachers necessary for joint revision of the instructional process.

"... there is no easy way to assess whether or not a specific criterion-referenced test is appropriate in any given situation."

Selecting Appropriate Criterion-Referenced Tests

Unfortunately, there is no easy way to assess whether or not a specific criterion-referenced test is appropriate in any given situation. Numerous factors determine what type of criterion-referenced test would be useful; and an equally large number of factors influence the quality of any single test of that type. Wise selection of criterion-referenced tests necessitates an adequate awareness of the types of tests available and of the decisions test developers have had to make in designing a useful instrument for assessment.

In examining criterion-referenced tests, four features should stand out as common to all instruments eligible for the classification (Note 1). The first, and perhaps the most meaningful to teachers, is a behavioral statement of each skill assessed by the test. This usually appears as a list of behavioral objectives but is sometimes a more detailed description of a behavioral domain (Popham, 1975) or a checklist of tasks and subtasks. The second feature that criterion-referenced tests have in common is a means of communicating the acceptable standard of performance for each behavior or skill. This, the "criterion," might be individualized for each skill or it might be standard across all skills evaluated. The third essential feature is, of course, the test items themselves. Although usually of the same paper and pencil type found on norm-referenced tests, criterion-referenced assessment is conducive to less artificial ways of evaluating knowledge or skill, and these methods are sometimes employed. Regardless of format the test items should be selected or designed specifically to represent

the exact behaviors stated in the objectives. The fourth and final feature is some form of data reporting system. Although the format for this may vary considerably, it is critical that specific information on student performance be communicated to the teacher for each behavioral statement, skill, or objective. This allows the teacher to plan instruction around each student's individual strengths and weaknesses.

Variations in Criterion-Referenced Tests

Both beyond and within these four essential features, one will find considerable variation among commercially marketed criterion-referenced tests. Knowledge of the possible variations can be useful when critiquing a test for purchase and when interpreting the meaning of student performance.



“... one of the most useful additional components is a prescriptive guide or reference list of instructional materials.”

Additional Components

Many tests have one or more additional components which can greatly assist teachers in administering the tests and interpreting the results afterward. Some tests, for example, have included a survey or placement instrument which allows the teacher to estimate a student's level of functioning. This information is useful in selecting the appropriate portions of a criterion-referenced instrument to be given to specific students. Some tests include cassette tapes of the test instructions, thus freeing the teacher and allowing students to administer the tests themselves. This feature is especially helpful if the test is to be used frequently to monitor student progress but is a wasted expense if the students' disabilities prevent taking advantage of it. As this would be the case for most hearing-impaired students, their teachers should exercise caution when considering such instruments. Other criterion-referenced tests are structured to use computer assisted scoring and reporting services. Although definitely a time saver, such services tend to separate the teacher from the student's test performance. Tests which provide computer assisted scoring frequently include performance profiles for the entire class and/or school, in addition to the usual individual student profile of skills. These can be useful for grouping students for instruction. Recording instruments for developing a class profile are also sometimes included in other criterion-referenced tests. Perhaps one of the most useful additional components is a prescriptive guide or reference list of instructional materials. It is usually correlated by page number to the test's objectives and can be im-

mensely helpful in locating the resources necessary to provide individualized instruction on specific skills.

Differences in Test Construction

Much less obvious to teachers, but far more critical in determining whether or not a criterion-referenced test will ultimately be useful for planning and evaluating instruction, are the differences in tests, most of which arise from decisions made when the tests are constructed. Criterion-referenced tests vary considerably in the amount of content assessed, the specificity of the behaviors measured, the content of the objectives selected, the rigor of the criterion established as reflective of acceptable performance, the number and type of test items per behavioral objective, and the manner in which student performance is recorded and reported. Each of these is discussed briefly in an attempt to acquaint teachers with the wide range of variables influencing the construction of a criterion-referenced test and their implications for the test's utility.

1. *Amount of content evaluated.* Some criterion-referenced tests have been developed by commercial publishers to evaluate a wide spectrum of skills across all grade levels in a single curriculum content area. The *Fountain Valley Teacher Support System* (1972-74), for example, assesses 785 different behavioral objectives in math alone. By contrast, other tests are designed to merely sample the possible range of skills at any given level. The *Basic Arithmetic Skills Evaluation (BASE)* (1973-74) provides test items for a selection of approximately 20 skills per grade level. Even more narrow are those tests designed to assess

student performance for a single unit of instruction (capitalization or time-telling, for example) and those developed to measure functional ability on a single task like tooth-brushing. In selecting a criterion-referenced test it is important to match the test to classroom needs. If there is a wide range in student abilities, for example, a test which focuses on skills from a single-grade level would probably not be adequate. On the other hand, if a number of students are in need of intensive instruction on a relatively well defined set of skills, a test providing a narrow but in-depth assessment of a single skill may be ideal.

2. *Specificity of behaviors measured.* Although the intent of a criterion-referenced test is to provide objective information about a student's ability to perform a well-defined behavior, there is incredible variance of opinion over what constitutes a single learner behavior. For some criterion-referenced tests, a single behavior may be as specific as the "ability to correctly add $4+2$." Another test, however, might define a single behavior as "the ability to correctly add all single-digit combinations with sums of 12 or less." Most criterion-referenced tests actually provide teachers with information about student performance in a series of "behavior-domains" rather than isolated skills. There is a limit to the amount of information a teacher can effectively process for decision-making (Popham, 1975), and comprehensive tests with hundreds of isolated skills to be evaluated may stretch beyond this limit. Teachers should be cognizant, however, of the specificity they feel is desirable and select a test which corresponds with their own biases.

3. *Content of objectives selected.*

"...teachers should examine the objectives or behavioral statements closely and choose a test which corresponds with their perceptions of the critical skills ..."



Related to specificity of the behaviors measured is the issue of *which* behaviors are measured. To assume that two criterion-referenced tests in reading (or even in a specific performance area such as "structural analysis") require the student to demonstrate the same skills is erroneous. The variations in possible task demands for even a relatively simple behavioral domain are numerous and not to be considered equivalent. If selecting a criterion-referenced test, teachers should examine the objectives or behavioral statements closely and choose a test which corresponds with their perceptions of the critical skills for that content area. Some tests in mathematics, for example, evidence traces of the 1960's movement toward "New Math." Set theory and working with bases other than 10 are no longer generally taught at the elementary level, yet quite a number of

criterion-referenced tests include objectives and items to measure such skills. When interpreting the results of a criterion-referenced test, one should take care to refrain from generalizing the results of the tests to skills not actually measured.

4. *Type and level of criterion.* In constructing a criterion-referenced test, one faces two major decisions with respect to the criterion or standard of acceptable performance. First, what *type* of criterion is appropriate and, second, what *level* of criterion is appropriate. An informal survey of criterion-referenced tests reveals that many types of criteria have been used. Among those are: (a) percent of test items correct per objective, (b) rate (number of items correct per time period), (c) description of the desired outcome (the criteria for well-formed letters, for instance), (d) duration of responding (usually in the form of a time

limit), (e) response latency (the lag between a command to perform a skill and the student's behavior), and (f) the amount of teacher assistance to be allowed. The specific type of criterion used for any specific behavior should reflect the purpose of the test and the performance level that is meaningful for that skill.

It is perhaps ironic that the name-sake of a criterion-referenced test, i.e., the criterion level to which student performance is referenced, is probably the most arbitrarily selected feature of the entire test. At the present time there is no good research to guide test developers, whether they be commercial publishers or teachers, in the selection of an appropriate standard for acceptable performance. Numerous real-world factors and constraints influence the selection of a criterion level and teachers aware of these variables are better equipped to interpret the tests' results accurately. In assigning a criterion level to a behavioral statement or objective, test developers are influenced by (a) the specific nature of the skill, i.e., how critical is it for student survival; (b) the proficiency required for the skill to be useful, usually based on real world demands; (c) the proficiency required for the student to profit from later instruction; (d) the number of test items and the criterion type previously selected; and (e) the characteristics of the population for whom the test is designed, e.g., a slower oral reading rate would be appropriate for hearing impaired students.

5. *The number and characteristics of test items.* In designing a criterion-referenced test, the ultimate goal is to select or create test items which accurately reflect the behavior or skill which is being evaluated. The degree of success with which this is

**"... there is no completely satisfactory way
for test items to be selected or constructed ..."**

accomplished has a tremendous influence on the generalizability of the test results and on the validity and reliability of the test. Even a relatively specific objective has hundreds and perhaps thousands of test items which could be considered representative of the skill being assessed. For example, Klein and Kosecoff (1973) have illustrated that even the following objective, as specific as it may appear, could easily generate an item pool of more than a thousand items:

The student can compute the correct product of two single digit numerals greater than 0 where the maximum value of this product does not exceed 20. (p. 4)

There are 29 pairs of numbers which could qualify for inclusion in the above objective and at least 10 alternate formats for writing the items,

6

e.g., $6 \times 8 =$, $8, 6 \cdot 8 =$, etc. This results in a total of at least 290 combinations for this one objective. Additional modifications such as altering the sequence of numerals, using different formats (multiple choice vs. completion), changing the mode of presentation (written vs. oral vs. taped), and changing the mode of response all serve to increase the potential item pool from hundreds to thousands (Klein & Kosecoff, 1973). Each of the above alterations can have an effect on the difficulty of the test item as well as on the degree to which the test item matches the presentation of the skill in the student's curriculum materials.

Of more concern to teachers, perhaps, is the number of test items selected to represent a particular objective or behavioral domain. Although the usual practice seems to be 3-5 test items per objective, some

criterion-referenced tests have as few as a single item for some skills and as many as 12 items for others. As there are no empirically based guidelines on the number of items needed to adequately represent a targeted behavior or skill, the decision is apparently determined more by such practical constraints as desired test length (Klein & Kosecoff, 1973). In short, there is no completely satisfactory way for test items to be selected or constructed, and teachers should examine any prospective purchase carefully.

6. *Reporting format.* Even the best constructed criterion-referenced test will be of limited assistance to teachers if student performance is not communicated in a way which is concise yet meaningful. Most criterion-referenced tests include some form of summary sheet or student profile with a listing of the behaviors or skills measured. There are differences, however, in the way a student's test results are recorded. The most common option is for each student's test performance to be compared with the established criterion for each skill and then evaluated either "pass" or "fail." This evaluation would be recorded, skill by skill, on the student's individual profile sheet, under columns labeled variously "know" or "don't know," "yes" or "no," "pass" or "fail," "reteach" or "proceed," "mastered" or "not mastered," etc. Another popular reporting format, and one frequently used in tests with computer assisted scoring, is a three-tiered communication of a student's level of mastery for each skill. The three levels are usually labeled with categories such as "mastered," "need to review," and "not mastered." Some criterion-referenced tests have been designed to report the number of test items an-

swered correctly or the number of errors per objective. This is especially common in tests with a standard number of test items per skill evaluated. Another frequently used reporting format is the simple checklist. Although similar in concept to reporting pass/fail, the checklist is more flexible if skills have been task-analyzed and reporting performance on all requisite subtasks is considered useful for the teacher. Checklists do not, of course, qualify as criterion-referenced tests if specific guidelines, i.e., criteria, for evaluating each skill are not described.

Some criterion-referenced tests also provide for recording the number of students passing each objective. Although not particularly useful for programming instruction, this information could be of great value in evaluating the success of past instructional procedures. Teachers selecting criterion-referenced tests should examine the manner in which student performance is communicated and assure themselves that the system used will provide the information necessary for efficient use in their classroom.

**Final Considerations in Selecting
and Using Criterion-Referenced
Tests**

It is probably obvious from the preceding section that the design and development of criterion-referenced tests necessitates making a variety of critical decisions about (a) objectives (which ones, how specific, how many); (b) criteria (what type, what level, how to report); and (c) test items (how many, how difficult, which ones, what format). The specific decisions made by a test developer have a profound influence on

"Criterion-referenced tests derive their strength from an ability to pinpoint specific skills requiring instruction and provide feedback on whether or not the instruction given was effective."

the ultimate utility of the test for classroom teachers and on the degree of confidence one can legitimately have in the test's results. Criterion-referenced tests are increasing in popularity because teachers are attracted by the idea of being able to make absolute interpretations about a student's performance, i.e., precise decisions concerning what a student can and cannot do. Unfortunately, absolute interpretations can be extremely dangerous, especially if accepted and acted upon without reference to the context from which they were derived. Although criterion-referenced tests hold great promise for planning and evaluating the instruction provided to handicapped students, they should not be accepted and used uncritically. As with any assessment instrument, the data derived from a criterion-referenced test should always be examined with full knowledge of the test's characteristics and possible weaknesses. Realizing that numerous decisions have had to be made throughout the test's design and construction, and interpreting test results in light of these decisions will help to prevent any abuse of the data derived from its administration.

A final recommendation to teachers is concerned more with test content than test quality. Although alluded to in previous sections, the concern is of adequate importance to warrant restating. In selecting and using a criterion-referenced test one is, in reality, making a commitment to a specific curriculum or a sequence of objectives for that particular content area. Because what one tests should be what one teaches, teachers need to make a decision to accept one of the following alternatives: (a) choose a criterion-referenced test and then develop or find an instructional

program to match, (b) choose an instructional program and then develop or find a criterion-referenced test to match (c) choose an instructional program with correlated tests provided, (d) choose a criterion-referenced test with correlated materials provided, or (e) develop a criterion-referenced test and an instructional program to match. Although some of these alternatives may be beyond the time or talents of classroom teachers, the importance of using criterion-referenced tests which measure precisely the same skills upon which there has been or will be instruction cannot be overemphasized. Criterion-referenced tests derive their strength from an ability to pinpoint specific skills requiring instruction and provide feedback on whether or not the instruction given was effective. Consequently, there is little value in assessing student performance on skills that will not be taught or teaching skills that will not be evaluated.

Reference Note

1. This does not mean, however, that all tests which claim to be criterion-referenced warrant the title. As with any movement growing in popularity, claiming an affinity is often profitable. Many tests designated "criterion-referenced" are actually norm-referenced tests with an accompanying list of objectives, or are objectives-based tests with no specified criterion for acceptable performance.

References

Basic education skills evaluation (BASE). Salem, Oreg.: Media Research Associates, 1973-74.

- Becker, W. C., & Engelmann, S. *Teaching 3: Evaluation of instruction*. Chicago: Science Research Associates, Inc., 1976.
- Block, J. H. Criterion-referenced measurements: Potential. *School Review*, 1971, 79, 289-298.
- Crehan, K. D. Item analysis for teacher-made mastery tests. *Journal of Educational Measurement*, 1974, 11, 255-263.
- Denham, C. H. Criterion-referenced, domain-referenced and norm-referenced measurement: A parallax view. *Educational Technology*, 1975, 15, 9-13.
- Fountain Valley teacher support system in mathematics*. Huntington Beach, CA.: Richard L. Zweig Associates, Inc., 1972-74.
- Glaser, R. Instruction technology and measurement of learning outcomes: Some questions. *American Psychologist*, 1963, 18, 519-521.
- Glaser, R., & Nitko, A. Measurement in learning and instruction. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, D.C.: American Council on Education, 1971, 652-670.
- Gronlund, N. E. *Preparing criterion referenced tests for classroom instruction*. New York: The Macmillan Company, 1973.
- Key math diagnostic arithmetic test*. Circle Pines, Minn.: American Guidance Service, 1971.
- Klein, S. P., & Kosecoff, J. Issues and procedures in the development of criterion-referenced tests. Princeton, N.J.: ERIC Clearinghouse on Tests, Measurement and Evaluation, 1973. (ERIC Document Reproduction Service No. ED 033 284)
- Olson, M. A. Criterion-referenced reading assessment in a large city school district. *Journal of Reading*

-
- Behavior*, 1976, 8, 387-395.
- Popham, W. J. Criterion referenced measurement. In W. J. Popham, *Educational evaluation*. Englewood Cliffs: Prentice-Hall, Inc., 1975.
- Popham, W. J., & Husek, T. R. Implications of criterion-referenced measurement. *Journal of Educational Measurement*, 1979, 6, 1-9.
- Sullivan, H. J. Developing effective objectives-based instruction. *Educational Technology*, July 1971, 55-57.